

**Stakeholder involvement in systematic reviews of  
healthcare interventions: reflections, reporting and  
recommendations**

**Rosalind Fay Wade**

**MPhil**

**University of York**

**Health Sciences**

**November 2022**

## **Abstract**

Systematic reviews use rigorous methods to identify, appraise and synthesise relevant studies to answer a research question. As well as being clearly reported and objective, it is important that they are clinically relevant and that the findings are applicable. Involving stakeholders in the review process, such as patients and healthcare professionals, may increase the relevance of the review and reduce the potential for research waste. Stakeholder involvement is valuable when interpreting systematic review results, putting them in context when drawing conclusions, improving applicability. Stakeholders are also well placed to help with dissemination and uptake of research findings.

Much of the existing literature on stakeholder involvement in healthcare research has focused on patient and public involvement in primary research. Literature on broader stakeholder involvement in systematic reviews is relatively scarce, although guidance has recently become available through the Cochrane online learning resource 'Involving People', developed using the evidence base identified by the ACTIVE (Authors and Consumers Together Impacting on eVidence) project.

This thesis describes and critically appraises the stakeholder involvement methods that I have developed over a range of review topics, including preventative, therapeutic and diagnostic healthcare interventions. For consistency and transparency, I have reported my methods using the ACTIVE framework, which has yet to be widely adopted by review authors. The (modified) ACTIVE summary table demonstrates the increasing level of stakeholder involvement and influence over the course of the projects, from contribution at specific review stages to control (or co-production). I have also presented a reflection/critical perspective, as recommended in the GRIPP2 (Guidance for Reporting Involvement of Patients and Public) checklist.

In order to supplement existing guidance, I have made specific recommendations for planning stakeholder involvement in systematic reviews, relating to budgeting/resources, recruitment, communication and practical considerations. These recommendations should be informative to researchers planning stakeholder involvement in systematic reviews.

## Table of contents

Abstract.....	2
Table of contents.....	3
List of tables .....	5
List of figures.....	6
Acknowledgements.....	7
Author’s declaration .....	8
Integrative chapter .....	11
Introduction and context .....	11
Aim and objectives .....	15
Development of my stakeholder involvement methods .....	16
Project 1: Antiembolism stockings for the prevention of deep vein thrombosis.....	16
Reflections/critical perspective.....	20
Project 2: Interventional management of hyperhidrosis .....	20
Reflections/critical perspective.....	25
Project 3: Ablative and non-surgical therapies for early hepatocellular carcinoma .....	26
Reflections/critical perspective.....	31
Project 4: Management of patients presenting to the Emergency Department with sudden onset severe headache .....	33
Reflections/critical perspective.....	37
Reporting stakeholder involvement in systematic reviews .....	40
Recommendations for planning stakeholder involvement in systematic reviews .....	45
Discussion.....	49
Limitations and areas for further research .....	49
References.....	51
Appendices .....	56
Appendix 1: Example email exchange with clinical advisors for the antiembolism stockings project .....	58
Appendix 2: Notes from the hyperhidrosis patient advisory group meeting.....	60
Appendix 3: Notes from the hyperhidrosis end-of-project workshop .....	68
Appendix 4: Collated notes from all hyperhidrosis stakeholder meetings (used to develop recommendations for further research).....	70
Appendix 5: Advertisement to recruit patient representatives to the project advisory group for the hepatocellular carcinoma project .....	75
Appendix 6: Information for patient representatives on the project advisory group for the hepatocellular carcinoma project .....	77

Appendix 7: Notes from the hepatocellular carcinoma project first advisory group meeting .....	80
Appendix 8: Notes from the hepatocellular carcinoma project second advisory group meeting .....	83
Appendix 9: Information for patient representatives on the project advisory group for the sudden onset severe headache project .....	87
Appendix 10: Project summary and agenda for end-of-project meetings with the project team and clinical advisory group members for the sudden onset severe headache project .....	90
Appendix 11: Notes from end-of-project meetings with the project team and clinical advisory group members for the sudden onset severe headache project .....	97
Appendix 12: Patient and Public Involvement section of the sudden onset severe headache RfPB report (following GRIPP2-SF guidance) .....	101
Appendix 13: Patient and clinician engagement section of the sudden onset severe headache full project report .....	102
Included papers .....	103

## **List of tables**

Table 1: Summary of involvement in each project using the ACTIVE framework	42
Table 2: Expertise and role of stakeholders in each project .....	44

## List of figures

Figure 1: Abstract for “Thigh length versus knee length antiembolism stockings for the prevention of deep vein thrombosis in postoperative surgical patients; a systematic review and network meta-analysis” <i>BMJ Open</i> , 2016.....	17
Figure 2: Abstract for “Systematic review of patient preference and adherence to the correct use of graduated compression stockings to prevent deep vein thrombosis in surgical patients” <i>Journal of Advanced Nursing</i> , 2017 .....	18
Figure 3: Abstract for “Interventional management of hyperhidrosis in secondary care: a systematic review” <i>British Journal of Dermatology</i> , 2018.....	21
Figure 4: Abstract for “Hyperhidrosis quality of life measures: review and patient perspective” <i>Journal of Dermatological Treatment</i> , 2019 .....	22
Figure 5: Patient and public involvement chapter of “Ablative and non-surgical therapies for early and very early hepatocellular carcinoma: a systematic review and network meta-analysis” <i>Health Technology Assessment</i> , 2023 .....	26
Figure 6: Abstract for “Management of patients presenting to the emergency department with sudden onset severe headache: systematic review of diagnostic accuracy studies” <i>Emergency Medicine Journal</i> , 2022 .....	34

## **Acknowledgements**

I would like to thank my supervisor, Professor Alison Eastwood, for advice throughout the preparation of this thesis and over the past two decades of my systematic reviewing career. I would also like to thank Mr Richard Morley, Ms Olivia Magwood and Dr Sarah Knowles for providing useful comments on an earlier draft. I am grateful to the co-authors of my publications, especially Dr Nerys Woolacott who I worked with on many diverse reviews over the years. I am also grateful to the patients and healthcare professionals who were involved in each of the projects for their input and for inspiring me to develop my methods of stakeholder involvement throughout these reviews. Finally, I would like to thank my friends and my family Jon, Charlie and Joe for their support.

## Author's declaration

The six papers included in this thesis are listed below, along with details of my contribution to each. The integrative chapter that links the papers is solely my own work. I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

Signed: 

Dated: 2 January 2024

Paper 1: Wade R, Paton F, Rice S, Stansby G, Millner P, Flavell H, Fox D, Woolacott N. Thigh length versus knee length antiembolism stockings for the prevention of deep vein thrombosis in postoperative surgical patients; a systematic review and network meta-analysis. *BMJ Open*. 2016;6:e009456. doi:10.1136/bmjopen-2015-009456.

*Contribution of the candidate: developed the protocol, recruited and spoke with the patient advisor, corresponded and met with clinical advisors, led the systematic review work (study selection, data extraction, validity assessment and synthesis of studies), prepared the manuscript and subsequent revisions.*

Signed: 

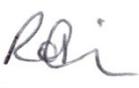
Signed: 

Rosalind Wade

Nerys Woolacott

Paper 2: Wade R, Paton F, Woolacott N. Systematic review of patient preference and adherence to the correct use of graduated compression stockings to prevent deep vein thrombosis in surgical patients. *Journal of Advanced Nursing*. 2017;73(2):336-48. doi:10.1111/jan.13148.

*Contribution of the candidate: developed the protocol, recruited and spoke with the patient advisor, corresponded and met with clinical advisors, led the systematic review work (study selection, data extraction, validity assessment and synthesis of studies), prepared the manuscript and subsequent revisions.*

Signed: 

Signed: 

Rosalind Wade

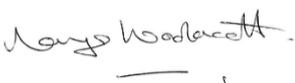
Nerys Woolacott

Paper 3: Wade R, Llewellyn A, Jones-Diette J, Wright K, Rice S, Layton AM, Levell NJ, Stansby G, Craig D, Woolacott N. Interventional management of hyperhidrosis in secondary care: a systematic review. *British Journal of Dermatology*. 2018;179(3):599-608. doi:10.1111/bjd.16558.

*Contribution of the candidate: co-applicant on the research proposal, developed the protocol, corresponded and met with stakeholders, led the systematic review work (study selection, data extraction, validity assessment and synthesis of studies), prepared the manuscript and subsequent revisions.*

Signed: 

Rosalind Wade

Signed: 

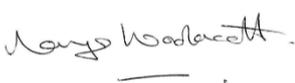
Nerys Woolacott

Paper 4: Wade R, Jones-Diette J, Wright K, Layton AM, Woolacott N. Hyperhidrosis quality of life measures: review and patient perspective. *Journal of Dermatological Treatment*. 2019;30:303-8. doi:10.1080/09546634.2018.1506080.

*Contribution of the candidate: co-applicant on the research proposal, developed the protocol, corresponded and met with stakeholders, contributed to the systematic review work (study selection, data extraction, validity assessment and synthesis of studies), prepared the manuscript and subsequent revisions.*

Signed: 

Rosalind Wade

Signed: 

Nerys Woolacott

Paper 5: Wade R, South E, Anwer S, Sharif-Hurst S, Harden M, Fulbright H, Hodgson R, Dias S, Simmonds M, Rowe I, Thornton P, Eastwood A. Ablative and non-surgical therapies for early and very early hepatocellular carcinoma: a systematic review and network meta-analysis. *Health Technology Assessment*. 2023;27(29). doi:10.3310/GK5221.

*Contribution of the candidate: co-applicant on the research proposal, developed the protocol, co-ordinated the patient involvement aspects of the project (recruited the patient advisors, held meetings with patients and acted as their contact on the research team), corresponded and met with clinical stakeholders, led the systematic review work (study selection, data extraction, validity assessment and synthesis of studies), prepared the systematic review elements of the report (including the patient and public involvement chapter) and subsequent revisions.*

Signed: 

Rosalind Wade

Signed: 

Alison Eastwood

Paper 6: Walton M, Hodgson R, Eastwood A, Harden M, Storey J, Hassan T, Randall M, Hassan A, Williams J, Wade R. Management of patients presenting to the emergency department with sudden onset severe headache: systematic review of diagnostic accuracy studies. *Emergency Medicine Journal*. 2022;39(11):818-25. doi:10.1136/emmermed-2021-211900.

*Contribution of the candidate: developed the research idea, led the funding proposal, developed the protocol, co-ordinated the stakeholder involvement aspects of the project (recruited, corresponded and met with stakeholders), led the systematic review work (study selection, data extraction, validity assessment and synthesis of studies), co-wrote the manuscript and subsequent revisions.*

Signed: 

Rosalind Wade

Signed: 

Matthew Walton

## **Integrative chapter**

### **Introduction and context**

Healthcare decisions should be informed by the best available research evidence. However, the vast amount of evidence available makes it difficult for busy clinicians and decision-makers to keep abreast of the latest, and most reliable, research about best practice.<sup>1</sup> Systematic reviews use rigorous methods to identify, evaluate and summarise the findings of relevant studies on a specific topic, making the evidence more accessible.<sup>1</sup> They provide valuable information on the effectiveness of healthcare interventions, maximising power, minimising bias and avoiding undue emphasis on individual study results.<sup>2</sup>

Systematic reviews involve defining a clear research question and implementing a comprehensive search strategy to identify all relevant studies to be systematically appraised and synthesised. Meta-analysis is often used to statistically synthesise data from several studies to produce a single quantitative estimate or summary effect size.<sup>3,4</sup>

It is imperative that the review question is meaningful for healthcare decision making; involving stakeholders is considered to increase the review's relevance, reducing the potential for research waste.<sup>5</sup> Stakeholders are defined as individuals or groups who are responsible for or affected by health- and healthcare-related decisions that can be informed by research evidence.<sup>6</sup> Stakeholders' contributions are valuable when interpreting the results of systematic reviews, putting them in context when drawing conclusions, improving applicability. Stakeholder involvement has been proposed as a way to enhance the usefulness and uptake of review findings.<sup>7</sup> Whilst the term 'stakeholder' is contentious in some settings, I have chosen to use this term, as it is widely understood in healthcare research; other terms include 'collaborators' and 'partners'.<sup>8</sup>

The 7Ps framework, developed in the USA, identifies seven key stakeholder groups in patient-centred outcomes research: patients and the public; providers; purchasers; payers; policymakers; product makers; and principal investigators.<sup>6</sup> The 7Ps framework has recently been expanded to include eleven groups: patients/consumers, caregivers, and patient groups; payers/funders of research; payers and purchasers of health services; publishers; policymakers; principal investigators; product makers; producers and commissioners of guidelines; program managers; healthcare providers; and the public.<sup>8</sup>

A mixed methods project, combining a literature review with key informant interviews, examined the benefits and challenges of stakeholder engagement in systematic reviews.<sup>9</sup> Expected benefits fell into six overarching domains: establishing credibility, anticipating controversy, ensuring transparency and accountability, improving relevance, enhancing quality, and increasing dissemination and uptake of review findings. Five overarching challenges were identified: (researcher and stakeholder) time and resources, researcher skills for stakeholder engagement, finding the right people, balancing multiple inputs, and understanding the most appropriate time to engage different types of stakeholders.<sup>9</sup>

Much of the existing literature on stakeholder involvement in healthcare research has focused on *patient and public involvement (PPI)*, which is now a requirement of many national funding bodies. The National Institute for Health and Care Research (NIHR) encourages researchers to involve patients, carers and the public throughout planning and delivering their research.<sup>10</sup> INVOLVE, an advisory group funded by the NIHR, was established in 1996 to support active public involvement in research.<sup>11</sup> Briefing notes for researchers were produced in 2003<sup>12</sup> and updated in 2012,<sup>13</sup> along with a supplement providing advice for researchers who design and carry out systematic reviews.<sup>14</sup> In 2019 the NIHR released a set of UK Standards for Public Involvement to improve the quality and consistency of PPI in health and care research.<sup>15</sup> The NIHR Centre for Engagement and Dissemination took over some INVOLVE functions in 2020.<sup>16</sup>

A systematic review assessing frameworks for supporting, evaluating and reporting PPI identified 65 frameworks with different provenances, intended purposes, strengths and limitations.<sup>17</sup> Primarily used by the groups who developed them, they were found to have limited transferability, suggesting that a single, off-the-shelf framework may be less useful than a set of evidence-based resources to be adapted for use. This demonstrates the size and diversity of the literature on PPI in research.

Published in 2011, the GRIPP (Guidance for Reporting Involvement of Patients and Public) checklist was the first attempt to develop robust guidance for reporting PPI activities.<sup>18</sup> In 2017, the GRIPP2 checklists were developed.<sup>19</sup> GRIPP2 long form includes 34 items and is suitable for studies where the main focus is PPI. GRIPP2 short form (GRIPP2-SF) includes five items and is suitable for studies where PPI is a secondary focus; this is the recommended format for reporting PPI in NIHR publications.<sup>20</sup> However, the checklists were not specifically developed

for *systematic reviews* and they provide guidance for reporting PPI, rather than wider *stakeholder* involvement.

Much of the impetus for *stakeholder* involvement in *systematic reviews* has come from the Cochrane Collaboration, a global organisation that collaborates to produce systematic reviews to support informed decision-making in healthcare.<sup>21</sup> The Cochrane handbook emphasises the importance of involving different stakeholders, including consumers, healthcare professionals, policy makers and funders, in order to increase the relevance of systematic reviews to a broad range of end users.<sup>5</sup>

A scoping review of reviews that reported stakeholder involvement found that the quality of reporting was generally very poor.<sup>22</sup> Only 10% of 291 included papers were judged to provide a comprehensive description of stakeholder involvement methods, half of which were methods papers. A range of stakeholders were included; 30% involved patients and/or carers, 41% involved other stakeholders (e.g., health professionals) and in 29% it was unclear who the stakeholders were. The level of stakeholder involvement varied from one-off to continuous.<sup>22</sup> This review informed the ACTIVE (Authors and Consumers Together Impacting on eVidencE) framework for describing the methods and approaches to stakeholder involvement in systematic reviews.<sup>7</sup> The ACTIVE continuum of involvement defines five levels of involvement: receiving, contributing, influencing, controlling and leading.<sup>7</sup> The term ‘co-production’ is increasingly being used to describe researchers, practitioners and the public working together, sharing power, ownership and responsibility.<sup>23-26</sup>

Cochrane Training developed an online learning resource, structured around the ACTIVE project, which provides guidance on practical issues to be considered when planning stakeholder involvement in systematic reviews.<sup>27</sup> These resources, alongside insights from various Cochrane groups and teams, also informed a six-step stakeholder engagement framework to support research groups in more broad areas of functioning, to meet a need for more stakeholder engagement support.<sup>28</sup> The Cochrane Co-Production Methods Group was launched in October 2023 to address evidence gaps to support the co-production of evidence syntheses.<sup>29</sup>

In August 2021 the Canadian Institutes of Health Research funded a 4-year project to develop guidance for multi-stakeholder engagement (MuSE) in health-related systematic reviews.<sup>30</sup> The project plans to develop equity-oriented guidance on methods for conducting, evaluating, and reporting engagement in evidence syntheses.<sup>8</sup> The MuSE Consortium has undertaken research on stakeholder

engagement in guideline development,<sup>31-35</sup> as well as providing practical guidance for involving stakeholders in health research.<sup>36,37</sup>

In summary, stakeholder involvement in systematic reviews can improve their relevance and applicability and potentially increase the uptake of their findings. Whilst there is a substantial body of literature on PPI in healthcare research, literature on *stakeholder* involvement in *systematic reviews* is relatively scarce. Guidance to support systematic review authors is available through a Cochrane online learning resource, developed using the evidence base identified by the ACTIVE project, including a scoping review of reviews reporting stakeholder involvement. However, as acknowledged by the authors, the scoping review relied upon review authors' reporting of methods, without additional clarification, so may lack sufficient detail to allow an in-depth appraisal of the methods.<sup>7</sup>

My stakeholder involvement methods developed over several years, alongside the developments described above. The review topics are diverse, including preventative, therapeutic and diagnostic interventions, and the level of stakeholder involvement and influence increased over the course of the projects, from contribution at specific review stages to control (co-production).

## **Aim and objectives**

The aim of this thesis is to clearly describe and critically appraise the methods I have developed for involving stakeholders in systematic reviews of healthcare interventions and make specific recommendations to supplement existing guidance.

The objectives are to: (1) describe my stakeholder involvement methods using the ACTIVE framework, which has yet to be widely adopted by review authors; (2) discuss the outcomes of stakeholder involvement in each project, and present a reflection/critical perspective, as recommended in the GRIPP2-SF checklist and, as a result of this critical reflection; (3) make specific recommendations for planning stakeholder involvement in systematic reviews, to supplement existing guidance.

## **Development of my stakeholder involvement methods**

As a researcher in a department that specialises in evidence synthesis, I have undertaken numerous systematic reviews in a range of healthcare topic areas. Each review was undertaken by a team which typically included a review manager with overall responsibility for the project, at least two reviewers, an information specialist/librarian and, depending on the objectives of the project, statistician(s) and/or health economist(s). For over two decades, the department has worked with advisory groups involving clinical experts to provide expertise of the specific health topic under review.<sup>38</sup> The involvement of patient experts is a more recent development within the department and requires a different approach to interaction with healthcare professionals, who are usually more familiar with research.

My initial experience of working with different stakeholder groups was in the antiembolism stockings project, which began in 2013. The methods used, particularly in terms of patient involvement, now appear rather limited and ‘tokenistic’ as this was an evolving field at the time. However, this initial experience helped me appreciate the benefit of incorporating different stakeholders’ perspectives in systematic reviews and prompted consideration of various issues, such as recruitment challenges and resource requirements.

### **Project 1: Antiembolism stockings for the prevention of deep vein thrombosis**

This project was commissioned by the NIHR Health Technology Assessment (HTA) programme to compare the relative effectiveness of thigh length versus knee length antiembolism stockings for the prevention of deep vein thrombosis in surgical patients. **Paper 1** (Figure 1) was published in a general medical journal in order to disseminate the findings to clinicians to inform future clinical practice.<sup>39</sup> A supplementary review was undertaken to assess patient preference and adherence to antiembolism stockings. **Paper 2** is a manuscript drafted specifically for a nursing audience summarising this review (Figure 2).<sup>40</sup>

**Figure 1: Abstract for “Thigh length versus knee length antiembolism stockings for the prevention of deep vein thrombosis in postoperative surgical patients; a systematic review and network meta-analysis” *BMJ Open*, 2016**

**Abstract**

**Objectives** To assess the clinical effectiveness of thigh length versus knee length antiembolism stockings for the prevention of deep vein thrombosis (DVT) in surgical patients.

**Design** Systematic review and meta-analysis using direct methods and network meta-analysis.

**Methods** Previous systematic reviews and electronic databases were searched to February 2014 for randomised controlled trials (RCTs) of thigh length or knee length antiembolism stockings in surgical patients. Study quality was assessed using the Cochrane Risk of Bias Tool. The primary outcome was incidence of DVT. Analysis of the DVT data was performed using ORs along with 95% CIs. The  $I^2$  statistic was used to quantify statistical heterogeneity.

**Results** 23 RCTs were included; there was substantial variation between the trials and many were poorly reported with an unclear risk of bias. Five RCTs directly comparing thigh length versus knee length stockings were pooled and the summary estimate of effect favouring thigh length stockings was not statistically significant (OR 1.48, 95% CI 0.80 to 2.73). 13 RCTs were included in the network meta-analysis; thigh length stockings with pharmacological prophylaxis were more effective than knee length stockings with pharmacological prophylaxis, but again results were not statistically significant (OR 1.76, 95% credible intervals 0.82 to 3.53).

**Conclusions** Thigh length stockings may be more effective than knee length stockings, but results did not reach statistical significance and the evidence base is weak. Further research to confirm this finding is unlikely to be worthwhile. While thigh length stockings appear to have superior efficacy, practical issues such as patient acceptability may prevent their wide use in clinical practice.

**Systematic review registration number** CRD42014007202.

**Figure 2: Abstract for “Systematic review of patient preference and adherence to the correct use of graduated compression stockings to prevent deep vein thrombosis in surgical patients” *Journal of Advanced Nursing*, 2017**

**Abstract**

**Aim:** The aim of this study was to explore patient preference and adherence to thigh and knee length graduated compression stockings for the prevention of deep vein thrombosis in surgical patients.

**Background:** Hospitalised patients are at risk of developing deep vein thrombosis. Mechanical methods of prophylaxis include compression stockings, available as knee or thigh length. Patient adherence to correct stocking use is of critical importance to their effectiveness.

**Design:** Systematic review of quantitative evidence.

**Data sources:** Eleven databases were searched from inception to 2013 for systematic reviews of compression stockings. Reviews were screened for relevant primary studies and update searches of eight electronic sources were undertaken (2010-2014).

**Review methods:** Randomised controlled trials and observational studies of surgical patients using compression stockings were quality assessed and data were extracted on patient adherence and preference. A narrative summary is presented.

**Results:** Nine randomised controlled trials and seven observational studies were included in the systematic review. There was substantial variation between studies in terms of patient characteristics, interventions and methods of outcome assessment.

**Conclusion:** Patient adherence was generally higher with knee length than thigh length stockings. However, the studies reflect patient adherence in a hospital setting only, where patients are observed by healthcare professionals; it is likely that adherence reduces once patients have been discharged from hospital. Patients preferred knee length stockings over thigh length stockings. In many clinical settings, any difference in efficacy between thigh length and knee length stockings may be rendered irrelevant by patient preference for and likely better adherence to knee length stockings.

**Keywords:** anti-embolism stocking; deep vein thrombosis; graduated compression stocking; literature review; nursing; patient adherence; patient preference; systematic review.

Who was involved?

In addition to the researchers in the project team, the project benefited from the expertise of an advisory group, including a vascular surgeon, an orthopaedic surgeon and an anticoagulant and thrombosis consultant nurse. Advice was also sought from a patient with experience of using antiembolism stockings after surgery.

How were stakeholders recruited?

Clinical experts were identified and invited to join the advisory group.

Unfortunately, an orthopaedic surgeon and two nurses initially contacted did not respond to emails. The orthopaedic surgeon who agreed to join the advisory group had personal experience of a deep vein thrombosis after surgery and was keen to

participate. The clinical advisors were not paid for their input, but were offered co-authorship of publications, if their contributions met journal authorship criteria.

Patient recruitment was informal; the patient was the mother of a colleague (not involved in this project) who made the introduction. I sent the patient a summary of the project, some questions for discussion and details of how we would pay her.

#### What was the mode of involvement?

Clinical advisors had continuous involvement through direct (telephone and email) interaction. The patient advisor had one-time involvement through direct (telephone) interaction.

#### Review stage and level of involvement

Healthcare professionals influenced the development of the protocol, data analysis, interpretation of review findings, review publication, and knowledge translation and impact. The vascular surgeon commented on the protocol; unfortunately, the other clinicians were not recruited until after the protocol was submitted. Clinical advisors were consulted via email and/or had telephone meetings with the project team when clinical questions arose requiring their expertise. No formal notes were taken at advisory group meetings, however, the text from one of the email exchanges with the clinical advisors is presented as Appendix 1. In addition to providing *ad hoc* advice, they also commented on the final report and co-authored the journal article.

The patient contributed during the protocol development stage. I offered to meet her separately from the main advisory group, if she preferred, which she did.

#### Outcomes of stakeholder involvement/impact

The clinical advisors provided advice when required, influencing the systematic review methods (e.g., their advice informed the inclusion of studies in the network meta-analysis, as shown in Appendix 1), and adding context when interpreting the review findings and discussing results in the journal article.

The involvement of a patient with experience of using antiembolism stockings at home gave insight into the practicalities of their use outside of a hospital setting, improving the applicability of the conclusions drawn.

### *Reflections/critical perspective*

This early experience of stakeholder involvement required minimal resources. Stakeholders were not paid for their contribution and correspondence was via telephone or email, incurring no travel costs. We had budgeted £75 for half a day of the patient's time (in line with rates suggested by INVOLVE), but she did not wish to be paid. The main resource was the time required to identify, recruit and consult with stakeholders, which took a few days of researcher time. Patient recruitment was informal and straightforward, arising from a casual discussion with a colleague; without this chance connection, additional time would have been required to identify a patient advisor. Recruitment of clinical advisors was more time consuming as the first three people approached did not respond. This highlights that initial contact with potential stakeholders should be made early to allow time for follow up emails and identifying additional suitable stakeholders.

We had hoped to collaborate with a local cardiology rehabilitation clinic, so that attending post-surgery patients could comment on our interpretation of the evidence, the economic modelling and discuss gaps in the evidence and proposed further research from a patient's perspective. Unfortunately, it was not possible to develop this collaboration (this aspect of the project was undertaken by a different member of the review team).

Reporting of stakeholder involvement was limited, prior to publication of the GRIPP2 checklist and the ACTIVE framework. There is now more scope to report stakeholder involvement more comprehensively.

### **Project 2: Interventional management of hyperhidrosis**

Building upon my earlier experience, the next two papers describe a project that benefited from the involvement of multiple patients and healthcare professionals, at various stages of the review process.

This review was undertaken as part of a larger project commissioned by the NIHR HTA programme to evaluate a range of interventions for hyperhidrosis. The objectives were to undertake a systematic review to estimate clinical effectiveness and inform key clinical parameters for a decision model, to develop a decision model to estimate cost-effectiveness, and to undertake a value of information analysis to help inform the design of future clinical studies.<sup>41</sup> Systematic reviews are used to inform the planning, design and conduct of new trials funded by the

NIHR HTA programme.<sup>42</sup> **Paper 3** describes the systematic review of clinical effectiveness evidence (Figure 3).<sup>43</sup> **Paper 4** describes an associated review of assessment tools used to measure health-related quality of life in hyperhidrosis research (Figure 4).<sup>44</sup>

**Figure 3: Abstract for “Interventional management of hyperhidrosis in secondary care: a systematic review” *British Journal of Dermatology*, 2018**

<p><b>Summary</b></p> <p><b>Background</b></p> <p>Hyperhidrosis is uncontrollable excessive sweating, which occurs at rest, regardless of temperature. The symptoms of hyperhidrosis can significantly affect quality of life.</p> <p><b>Objectives</b></p> <p>To undertake a systematic review of the clinical effectiveness and safety of treatments available in secondary care for the management of primary hyperhidrosis.</p> <p><b>Methods</b></p> <p>Fifteen databases (including trial registers) were searched to July 2016 to identify studies of secondary-care treatments for primary hyperhidrosis. For each intervention randomized controlled trials (RCTs) were included where available; where RCT evidence was lacking, nonrandomized trials or large prospective case series were included. Outcomes of interest included disease severity, sweat rate, quality of life, patient satisfaction and adverse events. Trial quality was assessed using a modified version of the Cochrane Risk of Bias tool. Results were pooled in pairwise meta-analyses where appropriate, otherwise a narrative synthesis was presented.</p> <p><b>Results</b></p> <p>Fifty studies were included in the review: 32 RCTs, 17 nonrandomized trials and one case series. The studies varied in terms of population, intervention and methods of outcome assessment. Most studies were small, at high risk of bias and poorly reported. The interventions assessed were iontophoresis, botulinum toxin (BTX) injections, anticholinergic medications, curettage and newer energy-based technologies that damage the sweat gland.</p> <p><b>Conclusions</b></p> <p>The evidence for the effectiveness and safety of treatments for primary hyperhidrosis is limited overall, and few firm conclusions can be drawn. However, there is moderate-quality evidence to support the use of BTX for axillary hyperhidrosis. A trial comparing BTX with iontophoresis for palmar hyperhidrosis is warranted.</p>
---

**Figure 4: Abstract for “Hyperhidrosis quality of life measures: review and patient perspective” *Journal of Dermatological Treatment*, 2019**

**Abstract**

**Purpose:** To identify the tools that have been used to measure quality of life in hyperhidrosis research and obtain patient insight on commonly used tools.

**Methods:** Twelve databases were searched to identify studies that reported measuring quality of life or described a quality of life tool in the context of hyperhidrosis. Data on the use of the tools were tabulated and hyperhidrosis-specific and dermatology-specific measures were summarized. A workshop was held to obtain the patients' perspective on the most commonly used tools and the newly developed HidroQoL tool.

**Results:** One hundred and eighty-two studies were included in the review. Twenty-two quality of life tools were identified; two or more tools were often used in combination. The most commonly used tools were the Hyperhidrosis Disease Severity Scale, the Dermatology Quality of Life Index and the Hyperhidrosis Quality-of-Life Questionnaire. Patient advisors preferred the new HidroQoL tool, which was considered to be easy to complete and most relevant to hyperhidrosis patients.

**Conclusions:** There are several tools available for assessing quality of life in hyperhidrosis patients; disease specific measures are widely used and appear suitable. It is unclear which tool is the most reliable, although the HidroQoL tool was preferred by a small group of patient advisors.

**Keywords:** Hyperhidrosis; dermatology; quality of life; review.

#### Who was involved?

In addition to researchers, the project team included two dermatologists and a vascular surgeon, each with expertise in managing patients with hyperhidrosis. A specialist nurse and four hyperhidrosis patients also provided advice at specific stages of the review.

#### How were stakeholders recruited?

The vascular surgeon had worked with us previously and agreed to work with us again; this pre-existing connection reduced the amount of time required for recruitment and provision of background information, as he was familiar with the process of stakeholder involvement in a systematic review. One dermatologist who was initially approached was unable to help us. However, as hyperhidrosis is not a rare condition, we were able to identify and recruit other suitable dermatologists. In addition, a nurse specialising in hyperhidrosis, who founded and ran the Hyperhidrosis UK support group, was invited to be a co-applicant on the proposal. However, she had already agreed to work on a competing bid, so could not be a co-applicant, but agreed to be an advisor once our bid was successful, both in her capacity as a specialist nurse and her role within the support group.

Two hyperhidrosis patient advisors were recruited through the clinical practice of one of the dermatologists (based in Harrogate, where a range of treatment options is available). Four patients were recruited to attend an end-of-project workshop. Again, patients were recruited through the clinical practice of the locally based dermatologist and the workshop was held at Harrogate District Hospital. Alternative approaches of recruiting patients via the Hyperhidrosis UK support group and the other dermatologist were considered, but it was logistically easier to recruit patients locally, as a face-to-face workshop was preferred. The dermatologist confirmed with the local research governance group that ethical approval was not required for the workshop, as patients were advisors to the project, rather than research participants.

Systematic review evidence indicated that further research was required on treatments for hyperhidrosis of the hand, therefore, we attempted to recruit patients with hand hyperhidrosis, alongside patients with axillary hyperhidrosis (which is more common). We also planned to recruit both male and female patients and patients of different ages; factors which we considered might affect the impact of hyperhidrosis on their lives. The dermatologist identified suitable patients and invited them to participate in the workshop; two of the patients had axillary hyperhidrosis and two patients had hand and axillary hyperhidrosis, three patients were female, one was male, and patients' ages ranged from their 20s to their 50s.

#### What was the mode of involvement?

Clinical advisors had continuous involvement through direct (face-to-face, telephone and email) interaction. Patient advisors had one-time involvement (at two different stages of the review process) through direct (face-to-face) interaction.

#### Review stage and level of involvement

Clinical members of the project team were involved from the proposal development stage. They influenced the development of the review question, methods, protocol, provided advice throughout the project (contributing to the search strategy), helped interpret the review findings in order to formulate conclusions and recommendations for further research, commented on the final report and dissemination activities. One of the dermatologists presented a poster on the project at the British Association of Dermatologists Annual Meeting, which won 'best psychodermatology poster' prize.<sup>45</sup>

Patient stakeholders influenced the interpretation of review findings and received information for dissemination, but did not provide comments that influenced the dissemination materials. Two patient advisors attended an initial meeting to discuss specific treatments, perceived effects of these treatments (both beneficial and adverse) and add context from a patient's perspective. Appendix 2 presents the transcribed notes from the meeting.

Four patient advisors and one dermatologist attended a workshop, held towards the end of the project. Prior to the workshop, participants were sent an overview of the project, copies of four quality of life tools that had been used in hyperhidrosis research and a short list of questions about the tools and were asked to consider them in preparation for the workshop. At the workshop an overview of the project was presented, along with a summary of the findings from the clinical effectiveness review and cost-effectiveness model. Gaps in the evidence base were discussed and the patients and dermatologist gave their opinions on the treatments and future research. The review of quality-of-life tools was described and patients commented on the tools and discussed important outcomes. Appendix 3 presents the notes from the workshop.

The results of the clinical effectiveness review, the cost-effectiveness modelling and proposed recommendations for future research were discussed further via teleconference with the other dermatologist and the hyperhidrosis specialist nurse (unfortunately the vascular surgeon was unavailable). The notes from all the end-of-project discussions with clinical and patient stakeholders were collated (see Appendix 4) and incorporated into our interpretation of the research findings and the project's conclusions and research recommendations.

A lay summary of the report was produced for circulation to three hyperhidrosis support groups. Patients who attended the workshop and additional patients identified by dermatologist stakeholders were invited to comment on the lay summary to ensure that it was clear and informative. The feedback was that patients found the summary interesting, agreed with the review findings about the clinical effectiveness of different treatments and the importance of assessing quality of life, and were glad that research was being undertaken for this condition. No specific suggestions for modification were made, which I considered to reflect the clarity of the summary, since it was written with a lay audience in mind.

### Outcomes of stakeholder involvement/impact

During the early stages, clinical stakeholders helped develop the proposal and protocol, increasing relevance and usefulness. Towards the end of the project, clinical and patient stakeholders helped interpret the results of the reviews and the cost-effectiveness model, ensuring that the conclusions and recommendations were relevant and applicable (see Appendices 3 and 4).

### *Reflections/critical perspective*

The recruitment of patients through the clinical practice of one of the dermatologists was straightforward and enabled the selection of patients with specific characteristics. This was particularly informative, for example the effect of treatment impairing hand sensitivity was highlighted by one of the patients with hyperhidrosis of the hand. Whilst alternative approaches to patient recruitment were considered, it was logistically easier to recruit patients locally. However, this could have resulted in the selection of patients with similar views, further to information they had received during their treatment.

The specialist nurse we invited to be a co-applicant on our proposal was unable to undertake the role, due to conflicting commitments, but agreed to be an advisor. This highlights the importance of contacting potential collaborators early.

The involvement of a larger group of stakeholders in this project required additional resources, which were costed into the project proposal. Based on rates suggested by INVOLVE, patients were paid £75 for attending each meeting, including preparatory reading (half a day in total); two patients attended the initial meeting and four patients attended the workshop. Clinical team members were costed at 2-3% of their salary for the 12-month period of the project. The specialist nurse was costed at £1000 (based on a daily rate of £300) to attend two meetings, answer queries throughout the project and comment on the protocol and final report. Other resources included travel expenses for patients and researchers to attend two meetings at Harrogate District Hospital, and additional researcher time preparing for and attending meetings (estimated at approximately 20 days in total).

A limitation of Paper 3 is the lack of reporting of stakeholder involvement methods; the methods section merely states that clinical and patient advisors contributed to the interpretation of results. Paper 4 includes a separate section describing the methods of collecting the patients' perspective and summarising their comments on quality-of-life tools used in hyperhidrosis research.

### **Project 3: Ablative and non-surgical therapies for early hepatocellular carcinoma**

**Paper 5** is the final report of a project that included a systematic review of ablative and non-surgical therapies for patients with early hepatocellular carcinoma, published in the NIHR HTA Journals Library.<sup>46</sup> Figure 5 presents the ‘Patient and public involvement’ chapter of the report, following GRIPP2-SF guidance, as recommended by the NIHR.<sup>19</sup>

**Figure 5: Patient and public involvement chapter of “Ablative and non-surgical therapies for early and very early hepatocellular carcinoma: a systematic review and network meta-analysis” *Health Technology Assessment*, 2023**

## **Chapter 6 Patient and public involvement**

### **Aim**

The aim of patient and public involvement was to ensure that the patient’s perspective was captured at all stages, from protocol development through to interpreting the results of the project and drawing conclusions and recommendations for further research.

### **Methods**

A patient collaborator was recruited to the project at the proposal writing stage via ‘Involvement@York’, the patient and public involvement network at the University of York. The patient collaborator attended all advisory group meetings and provided ongoing advice throughout the project. The patient collaborator was also consulted when producing materials in ‘plain English’, such as materials used when recruiting additional patients to the advisory group and the plain English summary section of the final report. The patient collaborator will be consulted during further dissemination activities.

Four additional patients were identified by our clinical advisors and recruited as members of the advisory group. With help from the patient collaborator, a lay summary of the project was produced describing the project, the role of advisory group members and details of how patients would be compensated for their time. This was circulated to patients who had expressed an interest in being a member of the advisory group. Patients were also provided with a lay summary of the different interventions included in the systematic review.

One member of the project team (RW) was the main contact for all patient advisors and held individual meetings with patients at the protocol development stage. During this initial meeting, patients were given background information to the project and a rudimentary description of the protocol and were asked for their comments, specifically whether any patient-relevant outcomes or aspects of treatment were missing from the protocol. Owing to the COVID-19 pandemic, all advisory group meetings were held via the Zoom™ online videoconferencing platform (Zoom Video Communications, San Jose, CA, USA), rather than in person. Patients were invited to attend the next advisory group meeting and the end-of-project workshop (see *Workshop*). Patients were also asked to comment on the final report.

**Figure 5 continued: Patient and public involvement chapter of “Ablative and non-surgical therapies for early and very early hepatocellular carcinoma: a systematic review and network meta-analysis” *Health Technology Assessment*, 2023**

### **Results**

All four patients were available at the beginning of the project to advise on the protocol. The patient collaborator and three of the patient advisory group members attended the second advisory group meeting held midway through the project to discuss the interim findings, prioritise interventions for further review and prioritise the most relevant patient outcomes. Patients provided helpful information about the outcomes most important to them, such as length of hospital stay and disruption to life (interventions requiring multiple appointments or repeat treatments) and level of pain involved. Non-recurrence of disease was another important outcome to patients. The patient collaborator and two patient advisory group members attended one of the end-of-project workshops. Unfortunately the other two patients were unavailable around the time of the workshops; in view of the reasons for their lack of availability, they were not pursued to attend at a different time. Patients were surprised by the lack of data on patient preference and quality-of-life outcomes in the existing evidence base. The patient collaborator and two patients commented on the final report.

### **Discussion and conclusions**

The patient and public involvement aspect of the project highlighted the outcomes most important to patients, which informed the development of the data extraction form. Their views added context to the review findings and their input was valuable when drawing conclusions and making recommendations for further research. The initial meeting with patients was informative to help the researchers understand the experience of patients, their concerns and preferences.

### **Reflective/critical perspective**

Patient involvement was a valuable part of this project, enabling researchers to understand important aspects of the different treatment options from a patient's perspective. One drawback was that meetings had to be held via the Zoom online videoconferencing platform, owing to the COVID-19 pandemic, which constrained the interactions with patients.

The feedback from patients was positive; they commented that information was presented clearly and that they found the meetings interesting and enjoyed being involved in the project.

### Who was involved?

In addition to researchers, the project team included a hepatologist and a patient collaborator. The project also benefited from the expertise of an advisory group, consisting of a clinical oncologist, a vascular and interventional radiologist, a hepatobiliary and general surgeon, a radiologist in diagnostic and interventional radiology, and four additional patients who had been treated for early hepatocellular carcinoma. I invited seven additional specialists to join the workshop, however, two of them did not respond and three did not attend, despite

agreeing to do so. Two additional clinicians attended; an interventional radiologist and a professor of hepatobiliary and transplant surgery.

#### How were stakeholders recruited?

The hepatologist project team member, who was a co-applicant on the proposal, had worked with members of the research team previously on a similar review topic. As described in Figure 5, the patient collaborator was recruited via the PPI network at the University of York (Involvement@York). Involvement@York is the central coordinating resource for PPI at the university, it serves as a recruitment 'hub', and provides information and support to academics and lay representatives. The patient collaborator had prior experience of involvement in research as a patient representative. However, she did not have direct experience of the specific condition under investigation, having received treatment for liver metastases, rather than primary hepatocellular carcinoma.

The clinical advisory group members were recruited through existing networks of the hepatologist; they were invited to join the advisory group and all of them responded promptly, agreeing to be involved. Our initial attempts to recruit patient advisory group members through Involvement@York were unsuccessful; members of the Involvement@York register (a group of approximately 75 patient and public members) and the Research and Development Unit at York Hospital were contacted. York Hospital responded that hepatocellular carcinoma patients from their region would be treated in Leeds, through the hepatologist project team member's multidisciplinary team. Therefore, patients were identified by one of the advisory group members, whose Personal Assistant forwarded an 'advertisement' to them, produced in collaboration with Involvement@York (presented in Appendix 5). Four patients emailed to say they would be happy to join the advisory group. The patient collaborator helped me to produce a document summarising the project and the role of the advisory group in more detail for the patient advisors (presented in Appendix 6). Owing to the nature of hepatocellular carcinoma, the hepatologist highlighted that unfortunately patient stakeholders may not all survive to the end of the project or may become too unwell to participate, therefore, we planned to recruit additional patients for the workshop (held towards the end of the project) if this was the case.

#### What was the mode of involvement?

Stakeholders had continuous involvement through direct (email and videoconferencing) interaction.

### Review stage and level of involvement

Clinical and patient stakeholders influenced the development of the review question, methods, protocol, provided advice throughout the project, influenced the interpretation of the review findings, the final report, and knowledge translation and impact. Clinical stakeholders also contributed to the search strategy, study selection, data extraction and data analysis.

The hepatologist project team member and patient collaborator worked closely with the research team from proposal development through to dissemination of the findings. The advisory group was established to provide advice throughout the project (responding to *ad hoc* queries from the research team), attend two advisory group meetings and an end-of-project workshop.

The first advisory group meeting was held to develop the research protocol. The second advisory group meeting was held just over half-way through the project to discuss the interim findings of the systematic review and prioritise interventions where RCT evidence was lacking but which were of particular interest and warranted targeted searching to identify non-randomised studies. The aim of the workshop was to discuss the findings of the systematic review and network meta-analysis, consider the feasibility of economic modelling and identify key priorities for further research and co-produce recommendations on which therapies, comparisons of therapies and trial outcomes should be prioritised for further research.

As we had been unable to recruit patients to the advisory group in time for the first meeting, I held individual meetings (via Zoom online videoconferencing platform) with the four patients to discuss the draft protocol, where I specifically asked about patient-relevant outcomes to be assessed in the systematic review and about which aspects of treatment were most important from a patient's perspective. Whilst it was more time consuming to hold individual meetings with patients, this one-to-one initial contact helped establish a rapport with patients and encouraged them to attend and contribute to future meetings.

For the second advisory group meeting patients were given the option of attending the main meeting or a separate meeting; all patients were happy to attend the full advisory group meeting, which was recorded using Zoom. Unfortunately, one patient was unavailable to attend the second meeting as he had to have a medical procedure, but expressed an interest in attending future meetings. Three clinicians

were also unable to attend the second advisory group meeting, although two of them provided comments after receiving notes from the meeting.

Owing to availability issues, two separate workshops were held. Two of the patients were unable to attend either workshop for personal reasons (ill health and family bereavement); in view of the reasons for their unavailability, I did not pursue alternative dates to meet with them. Stakeholders who attended the workshop were also asked to comment on the draft report. Specific sections where their comments would be particularly appreciated were highlighted, namely the plain English summary, scientific summary, the section describing the workshop discussions and the overall conclusions. Three clinical stakeholders and the patient collaborator sent comments on the draft report, primarily suggesting wording amendments and clarifying text relating to specific interventions, rather than the overall conclusions and recommendations, which were considered appropriate and reflected earlier discussions.

#### Outcomes of stakeholder involvement/impact

Stakeholders helped develop the proposal and protocol, increasing relevance and usefulness. Clinical stakeholders also provided additional advice when required, such as during study selection and data extraction stages (e.g., confirming the relevance of certain interventions/combinations of interventions to UK practice).

Notes from the first advisory group meeting are presented in Appendix 7, transcribed from a recording of the meeting. Clinical stakeholders identified additional interventions to be included in the review, which resulted in changes being made to the draft protocol and search strategies. Specific outcomes of interest were discussed, which resulted in amendments to the draft data extraction form. The patient collaborator requested an explanation of terms and methods prior to the workshop; clinical stakeholders helped to produce plain language summaries of relevant treatments.

During my first meeting with the patient advisors, they discussed the treatments they had received and when asked about outcomes of interest they considered that all relevant outcomes were already included in the protocol. Two issues were highlighted by patients that were beyond the scope of the research; one patient mentioned that a hospital information leaflet he received was not ideal for patients with poor eyesight (this had already been fed back to the treating clinician), two patients mentioned treatment delays encountered owing to the COVID-19

pandemic or referral to the specialist centre. These specific issues were not related to the review question; therefore, they were noted but did not result in changes to the protocol.

At the second advisory group meeting important interventions to be prioritised for further review and specific outcomes of interest were discussed. Patients expressed a preference for less invasive therapies that do not require multiple appointments or long hospital stays. They were disappointed at the lack of reporting of patient satisfaction outcomes in the trials that were identified by the systematic review. A list of interventions to be considered further was agreed; notes from the second advisory group meeting are presented in Appendix 8.

A document summarising the findings of the project was circulated to all workshop attendees in advance. At the workshop, patients again highlighted the lack of reporting of patient acceptability and quality-of-life outcomes as a limitation of the existing evidence base, resulting in our recommendation that these important outcomes should be addressed in future studies, alongside survival and recurrence outcomes. Clinical advice was informative for prioritising interventions of relevance to the NHS for further research, since many of the included studies were undertaken in East Asia where treatment approaches (and the underlying aetiology of liver disease) differ from the UK. Variation in treatment approaches among UK centres was also discussed. Detailed notes from the workshops (recorded using Zoom) are presented in Chapter 7 of the HTA report (Paper 5). The workshop discussions informed the conclusions of the report and recommendations for further research; stakeholder involvement helped ensure that they were applicable to UK clinical practice.

### *Reflections/critical perspective*

Multiple stakeholders were involved throughout this project, working together, sharing power, ownership and responsibility; the term ‘co-production’ is used to describe this level of involvement.<sup>23-26</sup>

Patient involvement at the proposal development stage required additional time investment at a point in the process with very strict timelines; corresponding with the PPI network and having initial discussions with the patient collaborator outlining the project and explaining research/medical terminology. However, the closer relationship with the patient collaborator and her continued involvement

throughout the project was informative and was particularly helpful when producing materials in plain English for other patient stakeholders.

Unfortunately, Involvement@York was unable to identify a patient with the specific condition under investigation, however, we were able to recruit a patient with a condition with similar treatment options. The patient collaborator had prior experience of involvement in research as a patient representative, which was helpful; the advantages of recruiting patients with research experience should be weighed-up against the advantages of recruiting patients with the specific condition under assessment. Additional patients were recruited to the advisory group through clinical stakeholders, as our initial attempts to recruit patients through Involvement@York were unsuccessful. Unfortunately, this meant that patients were not recruited in time for the first advisory group meeting.

Whilst all clinicians approached about joining the advisory group agreed promptly, there were problems recruiting additional experts to the workshop, with two clinicians not responding to emails and three clinicians not attending the workshop, despite having agreed to do so. In addition, two of the patient advisory group members were unable to attend. However, three patients and six clinicians attended, therefore, several stakeholders were able to contribute, discussing the findings of the project and co-producing recommendations for further research.

Stakeholder involvement was costed into the project proposal. The hepatologist's contribution was costed at 2% of his salary for 12 months. The patient collaborator was paid at a rate of £150 per day, in line with INVOLVE rates. Clinical members of the advisory group were not paid for their time but were acknowledged in the final report. Patient members of the advisory group were offered £75 per half day meeting/workshop and were acknowledged in the final report (after confirming that they were happy to be named). Two patients accepted payment, whilst two patients did not wish to be paid. The recruitment and involvement of a larger group of stakeholders, and attendance of several researchers at multiple advisory group meetings, required more researcher time resource than my previous projects: estimated at approximately 40 days in total. This is a significant time commitment, which is difficult to estimate at the beginning of a project; there were unforeseen time-consuming issues, such as problems recruiting patient stakeholders, and having to hold two separate workshops to accommodate the availability of multiple stakeholders.

Costs of transport, refreshments and room hire were included in the proposal (approximately £1700); the transport costs were estimated based on train fares for multiple stakeholders attending three meetings. These funds were not used, as restrictions imposed during the COVID-19 pandemic meant that all meetings had to be held via Zoom. However, despite the lack of face-to-face meetings, the positive feedback received from patients was encouraging and is hopefully a reflection, in part, of the rapport I developed with them.

PPI methods were reported following GRIPP2-SF guidance, as recommended by the NIHR.<sup>20</sup>

#### **Project 4: Management of patients presenting to the Emergency Department with sudden onset severe headache**

This project developed from an original question proposed by a consultant in acute internal medicine at Leeds Teaching Hospitals NHS Trust, during an informal discussion about research gaps in the area of emergency medicine. I undertook initial background literature searching and investigated potential funding streams. **Paper 6** describes a systematic review assessing the management of sudden onset severe headache patients presenting to the Emergency Department, funded by the NIHR Research for Patient Benefit Programme (RfPB) (Figure 6).<sup>47</sup>

**Figure 6: Abstract for “Management of patients presenting to the emergency department with sudden onset severe headache: systematic review of diagnostic accuracy studies” *Emergency Medicine Journal*, 2022**

**Abstract**

**Objective** Advances in imaging technologies have precipitated uncertainty and inconsistency in the management of neurologically intact patients presenting to the Emergency Department (ED) with non-traumatic sudden onset severe headache with a clinical suspicion of subarachnoid haemorrhage (SAH). The objective of this systematic review was to evaluate diagnostic strategies in these patients.

**Methods** Studies assessing any decision rule or diagnostic test for evaluating neurologically intact adults with a severe headache, reaching maximum intensity within 1 hour, were eligible. Eighteen databases (including MEDLINE and Embase) were searched. Quality was assessed using QUADAS-2. Where appropriate, hierarchical bivariate meta-analysis was used to synthesise diagnostic accuracy results.

**Results** Thirty-seven studies were included. Eight studies assessing the Ottawa SAH clinical decision rule were pooled; sensitivity 99.5% (95% CI 90.8 to 100), specificity 24% (95% CI 15.5 to 34.4). Four studies assessing CT within 6 hours of headache onset were pooled; sensitivity 98.7% (95% CI 96.5 to 100), specificity 100% (95% CI 99.7 to 100). The sensitivity of CT beyond 6 hours was considerably lower ( $\leq 90\%$ ; 2 studies). Three studies assessing lumbar puncture (LP; spectrophotometric analysis) following negative CT were pooled; sensitivity 100% (95% CI 100 to 100), specificity 95% (95% CI 86.0 to 98.5).

**Conclusion** The Ottawa SAH Rule rules out further investigation in only a small proportion of patients. CT undertaken within 6 hours (with expertise of a neuroradiologist or radiologist who routinely interprets brain images) is highly accurate and likely to be sufficient to rule out SAH; CT beyond 6 hours is much less sensitive. The CT–LP pathway is highly sensitive for detecting SAH and some alternative diagnoses, although LP results in some false positive results.

Who was involved?

In addition to researchers, the project team included a consultant in acute internal medicine, two consultants in emergency medicine and a consultant neurologist. A patient who had presented to the Emergency Department with a sudden onset severe headache was recruited as a patient collaborator.

The project also benefited from the expertise of an advisory group, including additional specialists in emergency medicine, acute and general medicine, neurology and neuroradiology, an NHS commissioner, and three additional patients who had presented to the Emergency Department with a sudden onset severe headache.

How were stakeholders recruited?

The acute medicine consultant discussed the proposed project with colleagues in emergency medicine and neurology, who were all keen to be involved and were co-applicants on the proposal. The patient collaborator (also a co-applicant) was identified by one of the clinical team members, I met with him to discuss the project and patient collaborator role, although no formal notes were taken. As a senior nurse in an Emergency Department, he understood acute medicine services

and diagnostic pathways, which was helpful and reduced the requirement for explanatory background material. However, he was not directly involved in assessing acute headache patients for suspected subarachnoid haemorrhage (SAH).

The clinical advisory group members were identified by the clinical members of the project team. They were selected to represent a broader range of perspectives; clinicians had expertise in different aspects of the management of headache patients (e.g., neuroradiology), or were based at smaller hospital trusts, with more limited access to neuroradiology expertise. The clinical members of the project team were all from Leeds Teaching Hospitals NHS Trust, which is one of the largest trusts in the UK.

Three additional patients were recruited early in the project with the help of a research nurse in the Emergency Department at Leeds General Infirmary, upon discharge from the headache pathway. Patients who agreed to join the advisory group were sent a summary of the project and their role (presented in Appendix 9).

#### What was the mode of involvement?

Clinical stakeholders and the patient collaborator had continuous involvement through direct (face-to-face, videoconferencing and email) interaction. The patient advisory group members had one-time involvement (at two different stages of the review process) through direct (face-to-face and telephone) interaction.

#### Review stage and level of involvement

The clinical and patient team members provided expertise throughout the project. The clinicians controlled the development of the review question. They influenced the development of the methods, protocol, provided advice throughout the project (contributing to the search strategy, selection of studies and data analysis), influenced the interpretation of the review findings, the final report, and knowledge translation and impact. The patient collaborator influenced the review question, contributed to the proposal, influenced the protocol, interpretation of the review findings, final report, and knowledge translation and impact.

Clinical members of the advisory group attended an initial meeting to discuss the draft protocol and care pathway for headache patients. The neuroradiologist also responded to email queries relating to specific computed tomography (CT) technologies during the study selection stage. Clinical stakeholders advised on

search terms; some 'headache' terms were considered too broad. Meetings were held either in person or via Zoom.

Separate meetings were held with each patient advisory group member to discuss the review protocol (including potential discussion points for the planned focus groups), important outcomes from a patient's perspective and to learn about patients' concerns and preferences regarding the care pathway. I met with one patient in a café, whilst the other two patients (both elderly ladies with health issues) preferred to have discussions over the telephone.

Towards the end of the project, I held additional meetings with stakeholders to discuss the systematic review findings. A project summary and agenda for the end-of-project meeting with the project team and clinical advisory group members is presented in Appendix 10. Individual telephone meetings were held with patient advisors.

The clinical and patient project team members also commented on the draft final report, journal manuscript and other dissemination activities (conference posters, summary report for clinicians and blog for headache patients published by The Migraine Trust).<sup>48</sup>

#### Outcomes of stakeholder involvement/impact

Discussions at the first advisory group meeting highlighted the variation in practice between hospital trusts and different departments. Clinical stakeholders helped develop the systematic review search strategy and inclusion criteria; it was agreed to include patients whose headache peaked within one hour and to exclude trauma patients, since they follow a different care pathway. The neuroradiologist suggested excluding studies older than 10-15 years, since scanners used at that time are not as accurate as those in current use. Inclusion criteria relating to outcomes of interest were considered to be comprehensive, although additional outcomes relating to costs (hospital bed days, admission/discharge rates) were suggested.

Initial meetings with patient advisors highlighted the complexities of managing patients who present to the Emergency Department with a sudden severe headache. Each of the patients had certain characteristics in their medical history which may have been associated with their headache symptoms; this helped us understand the decision problem more clearly and the difficulty managing patients in an emergency setting. The reasons for patients' preferences regarding undergoing

lumbar puncture after a negative CT scan highlighted that their fears and preconceptions were strong motivators.

Collated notes from stakeholder meetings held towards the end of the project are presented in Appendix 11 (produced from researchers' handwritten notes). The meetings helped with the presentation (format for presenting diagnostic accuracy results) and interpretation of the review findings and informed the conclusions of the review; stakeholders agreed that CT within six hours of headache onset appears to be sufficient to rule out subarachnoid haemorrhage, but it was agreed to emphasise the importance of access to neuroradiology expertise. Paper 6 highlighted the fact that risk tolerance of the patient and physician will continue to inform clinical practice. The paper also commented on the lack of evidence on the subgroup of patients who present to hospital several days after headache onset, as discussed at the meeting.

### *Reflections/critical perspective*

Stakeholder involvement helped researchers understand the clinical question and the complexities of managing patients in an emergency setting. Stakeholders helped develop the review question, increasing the relevance and usefulness of the review. Wider stakeholder involvement was informative when interpreting the results of the review and drawing conclusions, ensuring that they were applicable to different hospital settings. Stakeholders also commented on dissemination materials for clinicians and patients.

It was important to involve a range of clinical specialties involved in the management of patients presenting to the Emergency Department with headache. The risk tolerance of clinicians in different specialties varied, with emergency medicine clinicians (who see many headache patients) having a higher risk tolerance than neurologists (who only see those patients with a significant condition). We also recruited clinical advisory group members from different hospital settings in order to represent a broader range of perspectives. The recruitment of clinical stakeholders was straightforward; the research question was proposed by a clinician and the topic was considered to be very important. Recruiting patients who had experience of the specific care pathway under investigation was important, therefore, patients were recruited via a research nurse in the Emergency Department.

The four clinical co-applicants were each costed at 2% of their salaries for 12 months. The patient collaborator was costed at £1500; £150 per day for an estimated 10 days. The proposal included travel costs for team members (£500 to attend meetings with each other and the advisory group), the patient collaborator (£80 to attend approximately four meetings) and patient advisory group members (£120 to attend two meetings each), however, after initial meetings with stakeholders, COVID-19 restrictions meant all further meetings were held via telephone or Zoom. Researcher time for recruitment and consultation with stakeholders amounted to approximately 20 days.

Detailed information on patient involvement in the project was reported in the final submitted report, as the NIHR RfPB report template includes a section for reporting PPI, following GRIPP2-SF guidance (presented in Appendix 12). A full project report was produced and disseminated via the Centre for Reviews and Dissemination website,<sup>49</sup> which contains a section on patient and clinician engagement (presented in Appendix 13). However, the Emergency Medicine Journal submission guidelines simply request that authors provide a PPI statement in the methods section of the manuscript.<sup>50</sup> The strict word limit meant that full details could not be included.

A mixed methods approach had been planned for this project. Mixed methods research combines both quantitative and qualitative methods, providing greater insight and understanding.<sup>51</sup> The aim was to follow the systematic review with qualitative focus groups to explore patients' experiences of the management of headache and the acceptability of different care pathways identified in the review. We planned to integrate the findings of the qualitative research with the systematic review findings and discuss these with stakeholders in order to provide suggestions to policymakers and practitioners for improving patient management.

Unfortunately, the qualitative element of the project was severely impacted by the COVID-19 pandemic; changes to the patient pathway, reduced numbers of patients attending hospital and changing local and national restrictions affected patient recruitment. Therefore, it was not possible to undertake the focus groups or draw conclusions about the acceptability of different care pathways to patients. In the absence of the qualitative element of the project, the involvement of patient stakeholders provided some insight into patients' perspectives and preferences. However, stakeholder involvement should not be confused with qualitative research. Stakeholders are partners/advisors working with the research team to

improve the relevance and quality of the research, rather than research participants; the relationship and roles are different.<sup>52</sup> When patients (or healthcare professionals) are research participants, ethical approval should be sought well in advance of planned participant recruitment. Whilst ethical approval is not required for stakeholder involvement, potential harms to stakeholders should be carefully considered, particularly when vulnerable stakeholders are involved, as discussions about the research may have the potential to cause distress.

## Reporting stakeholder involvement in systematic reviews

In 2014, *The BMJ* introduced a policy requiring authors to report PPI activity within the methods section of manuscripts.<sup>53</sup> A study comparing the frequency of PPI reporting in research published in *The BMJ* before and after the introduction of the policy found that in the year before the policy, 0.5% of research papers reported PPI activity, whereas a year after the policy 11% reported PPI activity. The article suggested that the absence of information about PPI in papers is likely attributable to both a lack of reporting requirements and a lack of PPI activities.<sup>54</sup> A more widespread requirement for reporting stakeholder involvement in manuscripts is likely to encourage researchers to involve stakeholders. However, strict word limits imposed by journals make it difficult to report comprehensive details of stakeholder involvement activities. One possible solution is for journals to require authors to submit a supplementary appendix, specifically for reporting stakeholder involvement activities.

Clear, consistent reporting of stakeholder involvement activities facilitates understanding, evaluation and improvement in stakeholder involvement methods. A reporting framework or guideline provides a structure to increase transparency and standardisation, enabling the comparison of methods across reviews. The Equator (Enhancing the QUALity and Transparency Of health Research) Network define a reporting guideline as “A checklist, flow diagram, or structured text to guide authors in reporting a specific type of research, developed using explicit methodology.”<sup>55</sup>

In 2009, the PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) statement was published to improve the reporting of systematic reviews and meta-analyses.<sup>56,57</sup> The updated PRISMA 2020 statement specifies 27 items that should be reported.<sup>58</sup> Several journals state that PRISMA guidelines should be followed when reporting systematic reviews for consideration for publication.<sup>53,59-62</sup> However, none of the items in the guideline refer to stakeholder involvement. Consistent reporting of stakeholder involvement activities in systematic reviews could be encouraged by future updates of PRISMA including an item relating to the reporting of stakeholder involvement. The MuSE project, described in the introduction section of this thesis, plans to develop a guideline for reporting engagement in evidence syntheses as a PRISMA extension.<sup>8</sup>

A recent article describes a new taxonomy for defining the interests of stakeholders’ representatives in health research (in the context of guideline

development).<sup>63</sup> Potential conflicting interests of stakeholders (whether personal or of the organisation they represent) should be considered and declared in any project involving stakeholders, including systematic reviews.

I have combined the approaches of the ACTIVE framework<sup>7</sup> and GRIPP2-SF,<sup>19</sup> to clearly and consistently report and appraise my stakeholder involvement activities.

GRIPP2 is the first international guidance for reporting PPI activities in health and social care research.<sup>19</sup> It was not specifically developed for systematic reviews; the guidance for reporting PPI methods is less structured to allow its more general use. However, GRIPP2-SF prompted me to consider the outcomes of stakeholder involvement and to reflect upon and critically appraise my methods.

The ACTIVE framework provides a comprehensive structure for describing stakeholder involvement in systematic reviews.<sup>7</sup> The various review stages are defined, along with the different approaches, methods and levels of involvement (the ACTIVE continuum of involvement). Whilst I found the framework generally clear and comprehensible, judgements about the level of involvement required more consideration and felt more subjective. It is unclear how the 'receiving' level constitutes involvement, since it is described as not influencing the review process in any way. I prefer the term 'co-production' to 'control' when describing this collaborative level of involvement.

A summary of involvement in each project is presented in Table 1, using a modified version of the ACTIVE summary table; methods of patient and 'other stakeholder' involvement have been separated and the type of 'other stakeholder' is specified. In systematic reviews where different approaches are used for different stakeholder groups, this modification allows more accurate and complete reporting. For further clarity, Table 2 shows the expertise and role of the stakeholders in each project; this could be a helpful addition to the ACTIVE framework, when numerous stakeholders are involved in different roles within a project. In each of my projects, stakeholders who were members of the project team, rather than an advisory group, generally had more responsibility and influence on the project. The level of stakeholder involvement increased over the course of the projects, from contribution at specific review stages to control (or co-production). Only six of the systematic reviews that informed the ACTIVE framework involved patients and other stakeholders using a continuous or combined approach,<sup>64-69</sup> generally at fewer review stages than in my projects.

**Table 1: Summary of involvement in each project using the ACTIVE framework**

			What happened?		Stage and level of involvement												
Project	Who was involved?	How were they recruited?	Approach	Method	Develop question	Plan methods	Write protocol	Develop search	Run search	Select studies	Collect data	Assess risk of bias	Analyse data	Interpret findings	Write and Publish review	Knowledge translation	Top & tail?
Antiembolism stockings	Patients	Closed; invitation	One-time	Direct interaction			Ctb										No
	Other stakeholders: Healthcare professionals	Closed; invitation	Continuous	Direct interaction			Inf						Inf	Inf	Inf	Inf	Yes
Hyperhidrosis	Patients	Closed; invitation	One-time	Direct interaction										Inf		Rec	No
	Other stakeholders: Healthcare professionals	Closed; invitation	Continuous	Direct interaction	Inf	Inf	Inf	Ctb						Inf	Inf	Inf	Yes

Hepatocellular carcinoma	Patients	Closed; invitation and existing groups	Continuous	Direct interaction	Inf	Inf	Inf							Inf	Inf	Inf	Yes
	Other stakeholders: Healthcare professionals	Closed; invitation	Continuous	Direct interaction	Inf	Inf	Inf	Ctb		Ctb	Ctb		Ctb	Inf	Inf	Inf	No
Sudden onset severe headache	Patients	Closed; invitation	Combined	Direct interaction	Inf	Ctb	Inf							Inf	Inf	Inf	Yes
	Other stakeholders: Healthcare professionals	Closed; invitation	Continuous	Direct interaction	Con	Inf	Inf	Ctb		Ctb			Ctb	Inf	Inf	Inf	No

Blank (shaded) cells indicate that there was no stakeholder involvement at that review stage.

**Abbreviations:** Con: controlling; Inf: influencing; Ctb: contributing; Rec: receiving.

**Table 2: Expertise and role of stakeholders in each project**

	Project team member/co-applicant		Advisory group	
	Healthcare professional	Patient	Healthcare professional	Patient
Antiembolism stockings	-	-	Vascular surgeon Orthopaedic surgeon Consultant nurse	One patient with experience of using antiembolism stockings
Hyperhidrosis	Two dermatologists Vascular surgeon	-	Specialist nurse	Four hyperhidrosis patients
Hepatocellular carcinoma	Hepatologist	One patient with liver metastases	Clinical oncologist Vascular/interventional radiologist Hepatobiliary/general surgeon Diagnostic/interventional radiologist (additional interventional radiologist and hepatobiliary/transplant surgeon attended the workshop)	Four hepatocellular carcinoma patients
Sudden onset severe headache	Acute medicine consultant Two emergency medicine consultants Consultant neurologist	One sudden onset severe headache patient	Two emergency medicine consultants Emergency medicine registrar Acute and general medicine consultant Consultant neurologist Consultant neuroradiologist NHS commissioner	Three sudden onset severe headache patients

## **Recommendations for planning stakeholder involvement in systematic reviews**

This section presents specific recommendations, developed by reflecting on which aspects worked well and the difficulties encountered in the projects described, to supplement existing guidance in the Cochrane learning resource ‘Involving People’.<sup>27</sup> These recommendations address some of the identified challenges of involving stakeholders in systematic reviews.<sup>9</sup>

### *1) Make sufficient provision for stakeholder involvement and acknowledge contributions*

‘Involving People’ highlights that stakeholder involvement requires additional resources in the form of time and money; costs of the research staff who coordinate, support or facilitate involvement, costs of the activity that people are involved in and expenses of people involved.<sup>27</sup> Details relating to resources are presented in this thesis and exemplify the extensive resources required.

Clinical stakeholders with more extensive involvement in my projects were paid a proportion of their salary, based on estimates of the amount of time required. Clinicians who provided ‘one-time involvement’ were acknowledged or offered co-authorship of journal articles (if their contributions met journal authorship criteria) but were not offered payment.

Patient stakeholders were offered payment in accordance with INVOLVE recommended rates (along with any travel expenses),<sup>70</sup> although some patients did not wish to be paid. Updated guidance on payment for public involvement in health and care research was published in April 2023 to give direction relating to employment status and tax regulations and provide information and links to HMRC guidance.<sup>71</sup> Contributions of patient stakeholders should also be acknowledged in publications (along with declarations of potential conflicting interests), and consideration given to whether they meet journal authorship criteria; patients should be asked whether they would prefer to be named or remain anonymous.

Transport time/costs and refreshment costs were minimal in my projects, partly due to COVID-19 restrictions making it necessary for meetings to be held virtually. However, when face-to-face meetings are preferred, these costs can be substantial and should be calculated as accurately as possible (e.g., using advertised train fares) when planning a review and developing a proposal.

The additional researcher resource required for planning stakeholder involvement, recruiting stakeholders, preparing materials, organising and attending meetings and other activities is estimated and presented for each project, ranging from only a few days to around 40 days. This is a key component which should be carefully considered when planning stakeholder involvement and the amount of time required to undertake these activities should not be underestimated.

2) *Decide (early) who to involve, at what review stage, and how to recruit stakeholders (and have a back-up plan)*

Eleven key stakeholder groups are listed in the introduction section of this thesis, consideration should be given to which groups it may be appropriate to involve. Identification and recruitment of stakeholders should be considered at an early stage; stakeholder involvement during proposal and protocol development can help refine the research question and methods, increasing relevance. In a few of my projects, prospective clinical stakeholders did not respond to emails. Therefore, early initial contact with potential stakeholders is advisable to allow time for follow up emails and identifying other suitable stakeholders, if required.

Stakeholder involvement is valuable when interpreting review findings, putting them in context when drawing conclusions. In two of my projects the involvement of a wider group of clinical stakeholders at this stage was informative, incorporating the perspectives of clinicians from different settings. Stakeholders can strengthen dissemination activities, advising on how and where to present results and helping to produce publications and/or lay summaries.

In my experience, clinical stakeholders who had worked with us previously, or who were colleagues of clinicians involved in the project, were more likely to agree to be stakeholders than those not known to the project team. Therefore, it may be helpful to build up a network of clinicians who can be invited to be involved in future projects or recommend colleagues with the relevant expertise and make introductions. When undertaking a review in a new topic area, where the research team does not have existing connections, suitable clinicians can be identified by searching for relevant specialist groups and publications/guidelines and researching the members/authors.

Existing networks of clinicians were also helpful for identifying and recruiting patient stakeholders. Whilst this reduces researchers' workload, it could be argued that patients identified by clinical stakeholders are likely to reinforce the opinions of those clinicians, further to information they have received during their treatment.

Other sources available for recruiting patients include PPI networks, such as ‘Involvement@York’, charities or support groups for patients with the condition being assessed. PPI networks bring together patients (and the public) and researchers looking for lay representatives for specific research projects; patients are provided with information and support to help bring about meaningful, effective public involvement in research. However, PPI network members may not have experience of the specific condition being assessed. Therefore, it is important to weigh up the advantages of working with more experienced patient representatives who have support from a PPI network, or patients who have experience of the specific condition or interventions being evaluated. Some research topics may be more difficult to recruit patients to than others; our initial attempts to recruit hepatocellular carcinoma patients via Involvement@York were unsuccessful.

It may be helpful to recruit multiple patients with different characteristics, if certain characteristics are likely to affect their experiences or the effectiveness of the interventions under review, as in the hyperhidrosis project, where we recruited patients with hyperhidrosis affecting different body areas and patients of different ages and genders. Patients may be unavailable later in a project due to deteriorating health (as highlighted in our hepatocellular carcinoma project) or for other reasons. Therefore, consideration should be given to whether additional patients should be recruited during different stages of the project. It is important not to over-burden patients/caregivers, who may have more limited ‘spare’ time, whereas clinicians may consider involvement in research to be a part of their role.

### *3) Ensure communication is clear*

As highlighted in the ‘Involving People’ section on ‘Essentials for good practice’,<sup>27</sup> documentation provided to patients should be presented in plain language, e.g., lay summaries of the research methods and background to the review question. Appendix 5 presents an information sheet for prospective patients giving background information on the project and outlining the role of patient advisory group members. Appendices 6 and 9 give more specific information for patients who agreed to be advisors. These examples may be helpful to researchers planning to recruit patient stakeholders.

When communicating with stakeholders it can be beneficial to have one contact within the review team to help improve continuity. It is important to be approachable and attempt to establish a trusting and warm rapport. Whilst ethical

approval is not required, potential harms to stakeholders (particularly vulnerable groups) should be considered when communicating with stakeholders.

*4) Practicalities to consider when planning meetings*

Researchers need to be as flexible as possible when arranging meetings with stakeholders, some of which may need to be outside normal working hours to fit around clinicians' clinics and patients' working patterns. For the hepatocellular carcinoma project two workshops were held on different days, to allow as many stakeholders as possible to attend.

Using videoconferencing platforms, such as Zoom, are more widely accepted since the COVID-19 pandemic; using such technologies reduces travel time and means that meetings can be recorded for transcription. However, it is more difficult to develop a rapport with stakeholders without meeting in person, due to the lack of non-verbal cues and possible distractions, including technical issues.

It may be appropriate to offer patients the opportunity to meet separately, if they are not comfortable or confident meeting alongside a larger group of clinicians and/or researchers. In two of my projects patients expressed a preference to meet individually. However, it may be more difficult to balance multiple inputs when discussions are held out of context of wider group meetings, where different opinions can be explored and discussed.

## Discussion

This body of work presents my contribution to developing methods for involving stakeholders in systematic reviews of healthcare interventions. My methods are reported following the ACTIVE framework,<sup>7</sup> supplemented with specific examples of correspondence and notes from meetings with stakeholders. Each of the projects described builds upon the last; my expertise has developed over time and continues to inform the design of projects that I am involved in.

I have also presented a reflection/critical perspective, as recommended in the GRIPP2-SF checklist. Reflecting on which methods worked well, and the difficulties encountered, I have made specific recommendations for planning stakeholder involvement in systematic reviews, supplementing existing guidance.<sup>27</sup> Recommendations relate to making sufficient provision for stakeholder involvement, deciding who, when and how to recruit stakeholders, communication, and practical considerations. Whilst the stakeholders involved in my projects were limited to patients and healthcare professionals, my reflections and recommendations may be transferable to other stakeholder groups.

The papers presented in this thesis were published in medical journals, in order to inform clinical practice. The impact of my work has been acknowledged, with two papers being recognised by journal editors as amongst the most highly cited or downloaded articles in their journals,<sup>40,43</sup> and a third appearing on the journal's 'Most Read Articles' list for several months.<sup>47</sup>

### *Limitations and areas for further research*

A limitation of the work presented is the absence of people from minority groups amongst the patient stakeholders involved. I will try to increase inclusivity and diversity in future projects, e.g., by learning from the NIHR Ethnic Minority Research Inclusion group (e.g., ensuring that written and spoken materials are translated for people who don't have English as a first language and using resources to make information accessible to people who have difficulty reading).<sup>72</sup>

Another limitation is the lack of formal evaluation of the quality or impact of stakeholder involvement in the reviews. This is something that should be considered when planning reviews. The MuSE Consortium plans to develop guidance for evaluating the impact of stakeholder engagement in systematic reviews.<sup>30</sup> However, until such guidance is available, qualitative methods could be

used to elicit researchers' and stakeholders' views on the methods used and the perceived impact of stakeholder involvement. Both positive and negative aspects should be reflected upon, as outlined in GRIPP2-SF,<sup>19</sup> to inform and improve stakeholder involvement methods in future projects.

## References

1. CRD. *Systematic Reviews: CRD's guidance for undertaking reviews in health care*. Centre for Reviews and Dissemination; 2009.
2. Davey J, Turner RM, Clarke MJ, Higgins JPT. Characteristics of meta-analyses and their component studies in the Cochrane Database of Systematic Reviews: a cross-sectional, descriptive analysis. *BMC Medical Research Methodology*. 2011;11(1):160.
3. L'Abbe KA, Detsky AS, O'Rourke K. Meta-Analysis in Clinical Research. *Annals of Internal Medicine*. 1987;107(2):224-233.
4. Thacker SB. Meta-analysis: A Quantitative Approach to Research Integration. *JAMA*. 1988;259(11):1685-1689.
5. Higgins JPT, Thomas J, Chandler J, et al. (editors). *Cochrane Handbook for Systematic Reviews of Interventions version 6.2 (updated February 2021)*. Cochrane; 2021.
6. Concannon TW, Meissner P, Grunbaum JA, et al. A new taxonomy for stakeholder engagement in patient-centered outcomes research. *J Gen Intern Med*. Aug 2012;27(8):985-91. doi:10.1007/s11606-012-2037-1.
7. Pollock A, Campbell P, Struthers C, et al. Development of the ACTIVE framework to describe stakeholder involvement in systematic reviews. *Journal of Health Services Research & Policy*. 2019;24(4):245-255.
8. Tugwell P, Welch V, Magwood O, et al. Protocol for the development of guidance for collaborator and partner engagement in health care evidence syntheses. *Syst Rev*. Aug 2 2023;12(1):134. doi:10.1186/s13643-023-02279-1.
9. Cottrell E, Whitlock E, Kato E, et al. AHRQ Methods for Effective Health Care. *Defining the Benefits of Stakeholder Engagement in Systematic Reviews*. Agency for Healthcare Research and Quality (US); 2014.
10. NIHR. *Manage your project - Involve patients, carers and the public*. URL: <https://www.nihr.ac.uk/researchers/manage-your-funding/manage-your-project/#:~:text=Involve%20patients%2C%20carers%20and%20the%20public&text=Researchers%20are%20encouraged%20to%20evaluate,final%20report%20for%20their%20project> (accessed 14 April, 2023).
11. INVOLVE. URL: <https://www.invo.org.uk> (accessed 14 April, 2023).
12. INVOLVE. *Involving the public in NHS, public health and social care research: briefing notes for researchers*. 2003.
13. INVOLVE. *Briefing notes for researchers: involving the public in NHS, public health and social care research*. 2012.
14. INVOLVE. *Public involvement in systematic reviews: Supplement to the briefing notes for researchers*. 2012.
15. NIHR Centre for Engagement and Dissemination. *UK Standards for Public Involvement*. URL: <https://sites.google.com/nihr.ac.uk/pi-standards> (accessed 14 April, 2023).
16. NIHR. URL: <https://www.nihr.ac.uk/news/nihr-launches-new-centre-for-engagement-and-dissemination/24576> (accessed 14 April, 2023).
17. Greenhalgh T, Hinton L, Finlay T, et al. Frameworks for supporting patient and public involvement in research: Systematic review and co-design pilot. *Health Expectations*. 2019;22(4):785-801.
18. Staniszewska S, Brett J, Mockford C, Barber R. The GRIPP checklist: Strengthening the quality of patient and public involvement reporting in research. *International Journal of Technology Assessment in Health Care*. 2011;27(4):391-399.
19. Staniszewska S, Brett J, Simera I, et al. GRIPP2 reporting checklists: tools to improve reporting of patient and public involvement in research. *BMJ*. 2017;358:j3453.

20. NIHR. URL: <https://www.journalslibrary.nihr.ac.uk/information-for-authors/manuscript-preparation/reporting-patient-and-public-involvement.htm> (accessed 20 October, 2023).
21. Cochrane. URL: <https://www.cochrane.org/> (accessed 14 February, 2022).
22. Pollock A, Campbell P, Struthers C, et al. Stakeholder involvement in systematic reviews: a scoping review. *Systematic Reviews*. 2018;7(1):208.
23. Masterson D, Areskoug Josefsson K, Robert G et al. Mapping definitions of co-production and co-design in health and social care: A systematic scoping review providing lessons for the future. *Health Expectations*. 2022;25(3):902-913.
24. Smith B, Williams O, Bone L & the Moving Social Work Co-production Collective. Co-production: A resource to guide co-producing research in the sport, exercise, and health sciences. *Qualitative Research in Sport, Exercise and Health*. 2022:1-29.
25. Hickey G, Brearley S, Coldham T, et al. *Guidance on co-producing a research project*. INVOLVE; 2021. URL: <https://www.learningforinvolvement.org.uk/?opportunity=nihr-guidance-on-co-producing-a-research-project> (accessed 14 April 2023).
26. University of York. *Stronger Together: a guide for co-researchers working on co-produced research projects*. URL: <https://www.york.ac.uk/future-health/engagement/co-productionbooklet> (accessed 3 August, 2023).
27. Cochrane Training. *Involving People: A learning resource for systematic review authors*. URL: <https://training.cochrane.org/involving-people> (accessed 3 August, 2023).
28. Cochrane Training. *Six-step stakeholder engagement framework*. URL: <https://training.cochrane.org/online-learning/knowledge-translation/meaningful-partnerships/engaging-stakeholders> (accessed 28 July, 2023).
29. Cochrane Methods. *Cochrane Methods Co-Production*. URL: <https://methods.cochrane.org/co-production/news/launch-new-co-production-methods-group> (accessed 2 January, 2024).
30. Canadian Institutes of Health Research. *Stakeholder engagement in health care systematic reviews*. URL: [https://webapps.cihr-irsc.gc.ca/decisions/p/project\\_details.html?applId=445037&lang=en](https://webapps.cihr-irsc.gc.ca/decisions/p/project_details.html?applId=445037&lang=en) (accessed 3 August, 2023).
31. Petkovic J, Magwood O, Lytvyn L, et al. Key issues for stakeholder engagement in the development of health and healthcare guidelines. *Research Involvement and Engagement*. 2023;9(1):27. doi:10.1186/s40900-023-00433-6.
32. Petkovic J, Riddle A, Lytvyn L, et al. PROTOCOL: Guidance for stakeholder engagement in guideline development: A scoping review. *Campbell Systematic Reviews*. 2022;18(2):e1242. doi:10.1002/cl2.1242.
33. Magwood O, Riddle A, Petkovic J, et al. PROTOCOL: Barriers and facilitators to stakeholder engagement in health guideline development: A qualitative evidence synthesis. *Campbell Systematic Reviews*. 2022;18(2):e1237. doi:10.1002/cl2.1237.
34. Khabsa J, Petkovic J, Riddle A, et al. PROTOCOL: Conflict of interest issues when engaging stakeholders in health and healthcare guideline development: a systematic review. *Campbell Systematic Reviews*. 2022;18(2):e1232. doi:10.1002/cl2.1232.
35. Petkovic J, Riddle A, Akl EA, et al. Protocol for the development of guidance for stakeholder engagement in health and healthcare guideline development and implementation. *Systematic Reviews*. 2020;9(1):21. doi:10.1186/s13643-020-1272-5.
36. Concannon TW, Grant S, Welch V, et al. Practical Guidance for Involving Stakeholders in Health Research. *Journal of General Internal Medicine*. 2019;34(3):458-463. doi:10.1007/s11606-018-4738-6.
37. Parker R, Tomlinson E, Concannon TW, et al. Factors to Consider During Identification and Invitation of Individuals in a Multi-stakeholder Research

- Partnership. *Journal of General Internal Medicine*. 2022;37(16):4047-4053. doi:10.1007/s11606-022-07411-w.
38. Centre for Reviews and Dissemination. *Publications archive*. URL: <https://www.york.ac.uk/crd/publications/archive/> (accessed 22 November, 2023).
  39. Wade R, Paton F, Rice S, et al. Thigh length versus knee length antiembolism stockings for the prevention of deep vein thrombosis in postoperative surgical patients; a systematic review and network meta-analysis. *BMJ Open*. 2016;6(2):e009456.
  40. Wade R, Paton F, Woolacott N. Systematic review of patient preference and adherence to the correct use of graduated compression stockings to prevent deep vein thrombosis in surgical patients. *Journal of Advanced Nursing*. 2017;73(2):336-348.
  41. Wade R, Rice S, Llewellyn A, et al. Interventions for hyperhidrosis in secondary care: a systematic review and value-of-information analysis. *Health Technology Assessment*. December 2017;21(80). doi:10.3310/hta21800.
  42. Jones AP, Conroy E, Williamson PR, et al. The use of systematic reviews in the planning, design and conduct of randomised trials: a retrospective cohort of NIHR HTA funded trials. *BMC Medical Research Methodology*. 2013;13(50).
  43. Wade R, Llewellyn A, Jones-Diette J, et al. Interventional management of hyperhidrosis in secondary care: a systematic review. *British Journal of Dermatology*. 2018;179(3):599-608.
  44. Wade R, Jones-Diette J, Wright K, et al. Hyperhidrosis quality of life measures: review and patient perspective. *Journal of Dermatological Treatment*. 2019;30(3):303-308.
  45. Jones-Diette J, Wade R, Wright K, et al. *Hyperhidrosis Quality of Life Measures: Review and Patient Perspective*. URL: <https://www.york.ac.uk/media/crd/Hyperhidrosis-quality-of-life-measures.pdf> (accessed 25 March, 2022).
  46. Wade R, South E, Anwer S, et al. Ablative and non-surgical therapies for early and very early hepatocellular carcinoma: a systematic review and network meta-analysis. *Health Technology Assessment*. 2023;27(29). doi: 10.3310/GK5221.
  47. Walton M, Hodgson R, Eastwood A, et al. Management of patients presenting to the emergency department with sudden onset severe headache: systematic review of diagnostic accuracy studies. *Emergency Medicine Journal*. 2022;39(11):818-25 doi:10.1136/emmermed-2021-211900.
  48. Wade R. Managing sudden onset severe headache in the Emergency Department. 2021. URL: <https://migrainttrust.org/news/managing-sudden-onset-severe-headache-in-the-emergency-department/> (accessed 22 November 2023).
  49. Walton M, Hodgson R, Eastwood A, et al. *Management of sudden onset severe headache presenting to the Emergency Department: a systematic review and qualitative study*. 2021. URL: <https://www.york.ac.uk/media/crd/Sudden-onset-severe-headache-final-report.pdf> (accessed 28 July 2023).
  50. BMJ. *Emergency Medicine Journal - Submission guidelines*. URL: <https://emj.bmj.com/pages/authors/> (accessed 14 February, 2022).
  51. Creswell JW. *Research design: Qualitative, quantitative, and mixed methods approaches, 3rd ed*. Thousand Oaks, Ca, US: Sage Publications, Inc;2009.
  52. Hanley B, Staley K, Stewart D, Barber R. Qualitative research and patient and public involvement in health and social care research: What are the key differences? 2019. *NIHR Learning for Involvement*. URL: <https://www.learningforinvolvement.org.uk/content/resource/qualitative-research-and-patient-and-public-involvement-in-health-and-social-care-research-what-are-the-key-differences/> (accessed 16 October 2023).
  53. BMJ. URL: <https://www.bmj.com/about-bmj/resources-authors/article-types> (accessed 14 February, 2022).
  54. Price A, Schroter S, Snow R, et al. Frequency of reporting on patient and public involvement (PPI) in research studies published in a general medical journal: a descriptive study. *BMJ Open*. Mar 23 2018;8(3):e020452.

55. Equator Network. *What is a reporting guideline?* URL: <https://www.equator-network.org/about-us/what-is-a-reporting-guideline/> (accessed 16 October, 2023).
56. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA Statement for Reporting Systematic Reviews and Meta-Analyses of Studies That Evaluate Health Care Interventions: Explanation and Elaboration. *PLOS Medicine*. 2009;6(7):e1000100.
57. Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLOS Medicine*. 2009;6(7):e1000097.
58. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;372:n71.
59. NIHR. *Information for authors - Evidence synthesis and systematic reviews*. URL: <https://www.journalslibrary.nihr.ac.uk/information-for-authors/report-content/report-types/evidence-synthesis-and-systematic-reviews.htm> (accessed 14 April, 2023).
60. Lancet. *Preparing your manuscript*. URL: <https://www.thelancet.com/preparing-your-manuscript> (accessed 14 February, 2022).
61. JAMA. *Instructions for authors - reporting meta-analyses*. URL: <https://jamanetwork.com/journals/jama/pages/instructions-for-authors#SecReportingMeta-analyses> (accessed 28 March, 2022).
62. Annals of Internal Medicine. *Information for authors*. URL: <https://www.acpjournals.org/journal/aim/authors2> (accessed 14 February, 2022).
63. Akl EA, Khabisa J, Schünemann HJ, et al. A new taxonomy is proposed for defining the interests of stakeholders' representatives in health research: the case of guideline development. *Journal of Clinical Epidemiology*. 2023;160:147-150. doi:10.1016/j.jclinepi.2023.06.006.
64. Concannon TW, Fuster M, Saunders T, et al. A systematic review of stakeholder engagement in comparative effectiveness and patient-centered outcomes research. *J Gen Intern Med*. Dec 2014;29(12):1692-701. doi:10.1007/s11606-014-2878-x.
65. Coon JT, Gwernan-Jones R, Moore D, et al. End-user involvement in a systematic review of quantitative and qualitative research of non-pharmacological interventions for attention deficit hyperactivity disorder delivered in school settings: reflections on the impacts and challenges. *Health Expectations*. 2016;19(5):1084-1097. doi:10.1111/hex.12400.
66. Harris J, Croot L, Thompson J, Springett J. How stakeholder participation can contribute to systematic reviews of complex interventions. *Journal of Epidemiology and Community Health*. 2016;70(2):207-214. doi:10.1136/jech-2015-205701.
67. Liu JJ, Davidson E, Bhopal RS, et al. Adapting health promotion interventions to meet the needs of ethnic minority groups: mixed-methods evidence synthesis. *Health Technology Assessment*. 2012;16:44. doi:10.3310/hta16440.
68. Pollock A, Campbell P, Baer G, et al. User involvement in a Cochrane systematic review: using structured methods to enhance the clinical relevance, usefulness and usability of a systematic review update. *Systematic Reviews*. 2015;4(1):55. doi:10.1186/s13643-015-0023-5.
69. Rees R, Kavanagh J, Burchett H, et al. *HIV Health promotion and men who have sex with men (MSM): a systematic review of research relevant to the development and implementation of effective and appropriate interventions*. London: EPPI Centre, Social Science Research Unit, Institute of Education, University of London. 2004.
70. NIHR. *Payment guidance for researchers and professionals*. URL: <https://www.nihr.ac.uk/documents/payment-guidance-for-researchers-and-professionals/27392#payment-rates> (accessed 16 May, 2023).

71. NHS Health Research Authority. *Payment for public involvement in health and care research: a guide for organisations on employment status and tax*. URL: <https://www.hra.nhs.uk/planning-and-improving-research/best-practice/public-involvement/resources/payment-public-involvement-health-and-care-research-guide-organisations-employment-status-and-tax> (accessed 16 May, 2023).
72. NIHR Ethnic Minority Research Inclusion. *Ethnic Minority Research Inclusion Resources*. URL: <https://sites.google.com/nih.ac.uk/home/resources?authuser=0> (accessed 7 November 2023).

## **Appendices**

Appendix 1: Example email exchange with clinical advisors for the antiembolism stockings project

Appendix 2: Notes from the hyperhidrosis patient advisory group meeting

Appendix 3: Notes from the hyperhidrosis end-of-project workshop

Appendix 4: Collated notes from all hyperhidrosis stakeholder meetings (used to develop recommendations for further research)

Appendix 5: Advertisement to recruit patient representatives to the project advisory group for the hepatocellular carcinoma project

Appendix 6: Information for patient representatives on the project advisory group for the hepatocellular carcinoma project

Appendix 7: Notes from the hepatocellular carcinoma project first advisory group meeting

Appendix 8: Notes from the hepatocellular carcinoma project second advisory group meeting

Appendix 9: Information for patient representatives on the project advisory group for the sudden onset severe headache project

Appendix 10: Project summary and agenda for end-of-project meetings with the project team and clinical advisory group members for the sudden onset severe headache project

Appendix 11: Notes from end-of-project meetings with the project team and clinical advisory group members for the sudden onset severe headache project

Appendix 12: Patient and Public Involvement section of the sudden onset severe headache RfPB report (following GRIPP2-SF guidance)

Appendix 13: Patient and clinician engagement section of the sudden onset severe headache full project report

## Appendix 1: Example email exchange with clinical advisors for the antiembolism stockings project

Dear clinical advisor

We have been reviewing the evidence relating to knee length versus thigh length graduated compression stockings and have identified 23 RCTs assessing knee length or thigh length stockings. We are looking at the network of studies, to assess whether a network meta-analysis would be appropriate, and wondered whether we could combine studies of LMWH with studies of low dose heparin and/or studies of fondaparinux. Are these three different types of heparin interchangeable and their clinical effectiveness considered to be similar for the prevention of DVT?

Respondent 1: Yes I think they can be combined for this.

Respondent 2: I think it would be reasonable to combine LMWH/UFH/fondaparinux.

In addition, we wondered whether we could combine studies that assessed thigh length stockings alone with studies that assessed thigh length stockings alongside pharmacological prophylaxis, and studies of knee length stockings alone with studies that assessed knee length stockings alongside pharmacological prophylaxis. The only reason why we would be unable to combine these interventions is if the addition of pharmacological prophylaxis is likely to affect the *relative* effectiveness of knee length versus thigh length stockings - do you think this is likely?

Respondent 1: No – I DON'T THINK THIS IS LIKELY – but could you consider doing it separately as a sensitivity analysis?

Respondent 2: Whilst the addition of pharmacological prophylaxis is likely to affect the effectiveness of stockings I don't see why it would affect their relative effectiveness but don't see how we can know this without testing for it.

The studies we have identified are heterogeneous in terms of patient and surgical characteristics. Please could you indicate in the table below, which patient/study/surgical characteristics are likely to affect the *relative* effectiveness of knee length versus thigh length graduated compression stockings. We are aware that these characteristics are likely to affect a patient's risk of getting a DVT, but are interested in whether the characteristics would bias the study results in favour of either knee length or thigh length stockings.

*[One respondent completed the table]*

Characteristic	Likely to affect <i>relative</i> effectiveness of knee or thigh length GCS (yes or no)
Date of the study (i.e. 1970s, 1980s, 1990s, 2000s – reflecting differences in surgical methods/patient care over time)	No
Stocking applied to one leg versus both legs	No
Type of surgery (i.e. orthopaedic, abdominal, general, neurosurgery)	No (but thigh length may be difficult with some)

	procedures eg hip replacement)
Inclusion of patients with prior VTE	No
Inclusion of patients with active malignancy	No
Inclusion of patients with varicose veins	No
Inclusion of obese patients	No (see below)

Please let us know of any other characteristics likely to affect the *relative* effectiveness of knee length or thigh length GCS. Please could you send details of any supporting references, if applicable.

Respondent 1: The only issues are that in some patients the thigh length stockings are harder to wear so compliance may be lower in a thigh length group vs a knee length group. The 2 main things are leg shape with obesity tending to make it harder to fit the stocking properly and where there is surgery in the groin/hip.

Respondent 2: Likewise in the table other than the last item I don't see why any would affect their relative effectiveness. Obesity might as there might be an issue with one type fitting better than the other.

Respondent 3: My initial reaction would be should you be including studies from 30-40 years ago, as medical and nursing care has altered so much, that I am not sure it is relevant to today. I think that different chemical agents i.e. Low dose heparin and LMWH have shown different risk reductions so may affect outcome. Obesity would affect thigh length stockings.

Many thanks in anticipation of your advice.

Ros

## **Appendix 2: Notes from the hyperhidrosis patient advisory group meeting**

*Introduction:* Brief introduction to the objectives of the study and what we would like to gain from the patient advisory group meeting, i.e. contextualising our work and informing us as to any alternative analyses that we may wish to explore.

It will be made clear that the patients are fulfilling an advisory role to the project, advising on the appropriateness of the modelling assumptions. We will not be asking patients about their specific experiences, although of course their experiences will inform their advice.

The patients will be asked each question in turn. It may be that the patients cannot advise us on a particular question, in which case we simply move on to the next.

The patients will be encouraged to ask for clarifications at any point.

### Discussion points

#### **Background and progression of the condition:**

##### **Q.1**

(a) Hyperhidrosis emerges during adolescence.

Respondent 1: Can remember it emerging at around the age of 12 – adolescence would be a reasonable assumption.

Respondent 2: Slightly later than that.

(b) There may be seasonal, climatic and stress-related reasons for variation in the severity of hyperhidrosis, but the condition doesn't suddenly disappear or materially worsen during the ages of 18 to 65.

Respondent 1: Has been more or less the same – fairly constant.

Respondent 2: Has got worse as I have got older.

(c) The severity of hyperhidrosis or the impact on a person's life may reduce around the ages of 65/70.

No response to this assumption.

#### **Treatment:**

##### **Q.2**

(a) The model evaluates different sequences of treatments available or potentially available to patients on the NHS in secondary care; that is, referral to a specialist such as a dermatologist. These include iontophoresis (sponge or tray), oral medication, botox, curettage, endoscopic thoracic sympathectomy. We are evaluating all feasible sequences of these interventions such as botox followed by oral medication followed by iontophoresis. The only constraint is that curettage or ETS would never be offered before the other treatments. Is that a reasonable assumption?

Respondent 1: Yes, reasonable. Usually here patients are offered iontophoresis or the tablet form (oxybutynin) and if the tablet form doesn't work that's when you are usually offered botox. It can depend on what the patient prefers because sometimes it's a timing thing you can't actually come to the department to have iontophoresis so a tablet would be more likely, so it can just depend on the person really as well in terms of the treatment they have.

Respondent 2: No response to this assumption.

(b) Patients will try another treatment unless the current treatment isn't satisfactorily effective. A treatment may have a small benefit, but if it doesn't meet the patient's expectations and there is another treatment that could be tried then the patient will try it.

Respondent 1: Yes.

Respondent 2: Yes, definitely.

(c) After exhausting all treatments available to them, a patient will resort to a previous treatment that they tried which was only partially and not satisfactorily effective.

Respondent 1: Yes, definitely.

Respondent 2: If nothing else works, yes this is a reasonable assumption; anything that could help it slightly.

(d) A patient is most likely to retry a medication and then iontophoresis.

Respondent 1: You'd try anything.

Respondent 2: Yes, if nothing worked, you'd start again definitely.

*Interviewer follow-up: Maybe a different dose of medication, because that's the thing about medications you can vary the dose?*

Respondent 1: No response to this follow-up question.

Respondent 2: It depends what the side-effects are, some of the medications are quite strong and the side-effects do affect you more than having something else done which would make a difference.

(e) A patient only takes one treatment at any one time.

Respondent 1: Yes, reasonable.

Respondent 2: Yes, reasonable.

### **Q.3 Iontophoresis**

(a) A patient trials an iontophoresis device in a hospital for one month.

Respondent 1: You trial it, and then if it works the patient buys the machine and you can do it from home.

Respondent 2: No response to this assumption.

*Interviewer follow-up: So would you trial it for about a month?*

Respondent 1: Yes, for about a month and if the patient got on with it, great.

Respondent 2: No response to this follow-up question.

(b) Response to treatment (that is, treatment effectiveness) can be determined by one month.

Respondent 1: Never had it. For patients I know that have had it and it works for them within that month, they would buy a machine.

Respondent 2: Never had it, would be unable to answer.

(c) If the treatment is considered satisfactorily effective by one month, the same effectiveness is sustained indefinitely.

Respondent 1: No response to this assumption.

Respondent 2: No response to this assumption.

(d) Assuming the treatment is satisfactorily effective and there are no side effects serious enough to discontinue treatment, the patient purchases an iontophoresis device for the home to continue with that treatment.

Respondent 1: Yes, reasonable.

Respondent 2: No response to this assumption.

(e) A staff member is in attendance when a patient utilises an iontophoresis device in a hospital.

Respondent 1: Yes.

Respondent 2: No response to this assumption.

(f) If a patient continues to use iontophoresis at home, there would not be any planned follow-up visits with a specialist.

Respondent 1: Yes, reasonable.

Respondent 2: No response to this assumption.

(g) There may be side-effects sufficiently serious for a patient to stop treatment.

Respondent 1: Unsure.

Respondent 2: No response to this assumption.

#### **Q.4 Oral medication**

(a) The default assumption is that medication is taken at the recommended dose.

Respondent 1: Yes, reasonable.

Respondent 2: Yes, reasonable.

(b) There will be an alternative analysis where a patient takes medication only for key times of the week, say 4 or 5 times a week, in order to maintain treatment effectiveness. Reasonable alternative analysis?

Respondent 1: Yes, usually if you knew you were going out or something you'd probably take the tablet before you were going out rather than taking it in the morning so you knew that it works when you needed it.

Respondent 2: Side-effects also. Some people may not take it 5 times a day because of the side-effects, such as dry throat.

(c) Effectiveness is assumed to be temporary so in this scenario the patient only receives the benefit of the drug for a few hours.

Respondent 1: Yes, reasonable.

Respondent 2: I don't think it cures it no. Yes, reasonable.

(d) Response to treatment (that is, treatment effectiveness) can be determined by one month.

Respondent 1: It depends on how quickly it gets into your system I suppose.

Respondent 2: I'd think 3 months. So if you don't know after the first month if it's worked yet you should continue the course for the 3 months to see if it's really going to work. If it hasn't worked after 1 month it can be hard to say. If you stopped after 1 month, it might suddenly kick in after a couple of weeks and start working, you don't know. I think I was on it for about 3 months before nothing happened, but everyone's different I suppose.

(e) If the treatment is considered satisfactorily effective by one month, the same effectiveness is sustained indefinitely in the default case. In alternative scenarios, different declines in effectiveness over time will be tested.

Respondent 1: No response to this assumption.

Respondent 2: Yes, if it worked I'd assume it might be, but it's a tricky question to answer. Because then you'd permanently be on the medication wouldn't you?

(f) Assuming the treatment is satisfactorily effective and there are no side effects serious enough to discontinue treatment, the patient continues with that treatment indefinitely.

Respondent 1: Yes, but difficult to answer.

Respondent 2: Yes, but difficult to answer.

(g) There is a planned follow-up visit after the first three months with the dermatologist and then patients have monitoring visits with a GP every three months.

Respondent 1: Yes, they would have a follow-up visit with the dermatologist after three months. Patients may have an open appointment and if they feel it's not right for them they can come back to dermatology but yes, it would usually be the GP who would take over care.

Respondent 2: Yes, they would have a follow-up visit with the dermatologist after three months. Follow-up can depend on your side-effects, if you keep taking it all the time it may get worse.

*Interviewer follow-up: So I suppose what we're asking is if a GP would set up these regular follow-up visits? That's what we're currently assuming but maybe that's not appropriate. Maybe actually it's as and when the patient feels they need it.*

Respondent 1: No response to this follow-up question.

Respondent 2: Yes, I would think it's more on when the patient experiences side-effects and if the side-effects are getting worse they go back as opposed to going just in case they were to get worse. Even if the treatment worked but gave you side-effects, you would try something else rather than experiencing all of those side-effects.

*Comment: There is uncertainty around this issue. No real reason to think that the GP would arrange regular follow-up visits.*

(h) There may be side-effects sufficiently serious for a patient to stop treatment.

Respondent 1: Yes, this is reasonable. There's usually side-effects. Dry mouth and throat.

Respondent 2: Yes, this is reasonable. There are side-effects, dry mouth and throat. So even though it's helping one thing it can create problems elsewhere.

## **Q.5 Botox**

(a) Botox injections are repeated on average every 6 months.

Respondent 1: We do it yearly. If some patients are severe enough they would have it twice a year.

Respondent 2: No response to this assumption.

(b) The effectiveness of botox injections is fully sustained over the 6 months.

Respondent 1: It depends, everyone's different. Some people have it after 6 months and need it again after a couple of months and other people go the full year and they're ok.

Respondent 2: I think it comes back gradually too, it doesn't come back as bad as before you had the injections but if you didn't have any repeats at all it probably would do. It's not as severe though when it first starts to come back.

*Comment: There would appear to be a gradual reduction in effectiveness. Patients indicate that you know when you are ready again for a repeat injection.*

(c) There are no planned follow-up visits in between injections.

Respondent 1: You get a follow-up in the post to say that you need to come back and have your next lot of injections. This would usually be 1 or 2 months in advance.

Respondent 2: Mine didn't follow it up after the year, by which stage I was desperate for the injections again and they did follow it up after that.

*Interviewer follow-up: For the patients who are severe and require more than 1 injection in a year, presumably they don't receive the follow-up invitation in the post. How do they go about arranging their next appointment?*

Respondent 1: They would phone to say they would like to be seen again and they would have a consultation and the doctor would see if the patient should be seen again and that maybe that patient may require it twice yearly. For some patients that would work and for some people, botox just doesn't work for them at all.

Respondent 2: No response to this follow-up question.

*Comment: So in the case of changing the frequency of injections, you may have a consultation. However, for those with a regular treatment schedule there would be no need to meet with anyone – they would just get a reminder in the post a month in advance.*

*Interviewer follow-up: Is the length of effectiveness dependent on the severity of that patient's hyperhidrosis or is it more likely that some people just respond better to botox and it lasts longer in them?*

Respondent 1: I'm not sure, it's quite individual really.

Respondent 2: No response to this follow-up question

*Interviewer follow-up: I wonder would it be decided at the initial consultation as to whether that patient requires two injection a year etc.?*

Respondent 1: I would say the patient would have one treatment and if it hasn't worked the patient may ring and say that it's been two months and it hasn't worked, or longer and it hasn't worked, so they would call and the dermatologist would invite them back and offer them another treatment to see if it would work the second time.

Respondent 2: No response to this follow-up question.

*Interviewer follow-up: But if it works the default would be to offer a one year repeat?*

Respondent 1: Yes

Respondent 2: I think you can also put more in on one side of the body than the other which obviously helps with the effectiveness as well.

*Interviewer follow-up: The first visit, that's where most of the uncertainty is really as to whether or not it's going to be effective. So that first time, it's still just a reminder in the post? The assumption is that you will have it again.*

Respondent 1: Yes. When they come for treatment they will automatically get a reminder in the post after a year.

Respondent 2: I thought it was quite a new thing so I didn't know how frequently you would have it, whereas they're using it more now than they were back then.

(d) If the treatment is considered satisfactorily effective to repeat the injections, the same effectiveness is sustained indefinitely.

Respondent 1: No response to this assumption.

Respondent 2: You'd hope it would. If it was going to do that you would keep going back to have it done. Once it's done you don't need to think about medication or anything like that.

*Interviewer follow-up: Are the effects instant?*

Respondent 1: Sometimes it's after a 2-week period where I notice that I'm not sweating at all but then other times it can be immediately. 2 week maximum I would say before effectiveness is fully realised.

Respondent 2: It would be 1-2 weeks and it does make a really big difference. 2 weeks maximum, whereas with medication you could be on it for a couple of months before it was noticeable, but then again I suppose everyone's different.

(e) Other than the discomfort of the injections, there are no side effects sufficiently serious for a patient to avoid repeating the injections.

Respondent 1: No, I haven't experienced any. It is amazing really, life changing.

Respondent 2: No, none at all. There are no side-effects like the medication; the medication felt quite severe whereas this one didn't. It's a bit uncomfortable having it done but you would do anything.

*Interviewer follow-up: At Harrogate, is there anaesthetic while having the botox done?*

Respondent 1: No, just straight in.

Respondent 2: No.

*Interviewer follow-up: How many injections are there with botox?*

Respondent 1: 20 per arm. It's a tiny insulin syringe but it nips. But if it works it's worth it.

Respondent 2: Agreement with respondent 1.

*Interviewer follow-up: So does it take a while to actually do it?*

Respondent 1: About 10 minutes, unless you are a new-starter in which case you will need to get your consent form signed which may mean that it could take 15-20 minutes. However, after this first time it should be 10 minutes every time, very quick. It's 2 insulin syringes per arm; the same needle but 2 syringes. So 4 syringes in total. 10 doses in each syringe.

Respondent 2: About 10 minutes, very quick.

*Interviewer follow-up: Does a nurse generally carry out the procedure?*

Respondent 1: It's a doctor that does it here – dermatologist. There's talk of training up the staff nurses to do the procedure.

Respondent 2: Think that they are thinking of training up the staff nurses to do the procedure because of the appointment system – so that they could offer it to more people.

*Comment: Possible alternative scenario where we have a staff nurse do it rather than a dermatologist.*

*Interviewer follow-up: Is it a very uncomfortable procedure?*

Respondent 1: Agreement with respondent 2.

Respondent 2: Certain injections are. It can really pull. You're just desperate to have it done so it doesn't matter. It can make your eyes water because it's one injection after the other but you just do it.

*Interviewer follow-up: Uncomfortable rather than painful?*

Respondent 1: It is painful.

Respondent 2: No response to this follow-up question.

*Interviewer follow-up: There isn't some kind of local anaesthetic?*

Respondent 1: There is a local anaesthetic, we have a cream that numbs the skin.

Respondent 2: I don't think that these would reduce the discomfort.

## **Q.6 Endoscopic thoracic sympathectomy operation**

(a) If a patient experiences compensatory sweating then they will resort to a previous treatment that they tried which was only partially and not satisfactorily effective.

Respondent 1: No response to this assumption.

Respondent 2: You probably would, you'd go back and try it again.

(b) A patient is most likely to retry a medication and then iontophoresis.

Respondent 1: No response to this assumption.

Respondent 2: No response to this assumption.

### **Appendix 3: Notes from the hyperhidrosis end-of-project workshop**

Present: 4 patient advisors (two with HH of the axilla, two with HH of the hand and axilla), Alison Layton, Nerys Woolcott, Ros Wade, Steve Rice and Julija Stoniute

We went through the prepared PowerPoint presentation of clinical effectiveness results.

#### **Discussion of clinical review results**

Iontophoresis for hands: Not considered effective by patients and there are side effects. Also, relative to botox, much more patient time is spent on using iontophoresis, which was a consideration. One patient commented that she had received iontophoresis for the hands, which was not hugely effective, but had also received botox for the axilla, which really worked, therefore, she would be interested in botox for her hands. There was a general consensus that botox for axilla really worked. They liked the infrequent (annual administration), much preferred to frequent use of iontophoresis or application of creams. The patients agreed that future trials of treatments for hyperhidrosis of the axilla should compare against botox.

Botox for hands: Should be studied BUT pain on administration would have to be controlled and patients would need evidence and assurance that botox would not result in long term impairment of hand sensitivity.

Comparator treatments for a trial of botox for hands: Patients agreed iontophoresis would be an obvious comparator. Glycopyrrolate creams (works in some areas, not very effective in others, but 'ok') also used for HH of the hands.

Oral medications: Don't always work and side effects troublesome. Topical glycopyrrolate (to hand and other areas) didn't have the adverse effects of oral medications (dry eyes/mouth).

Curettage: patients would need assurance that it really was a 'one off' treatment and that it was effective – otherwise they would rather stick with annual botox (NB botox annual not 6 mths). Also significant concerns about scarring – would worry about actual appearance AND about being asked about scarring and not wanting to say they had HH.

From patients' perspective research into permanent treatments (like curettage) that reduced risk of scarring without reducing efficacy would be welcomed (e.g. laser, microwave, etc.).

Patients felt it wouldn't be worth doing further research on iontophoresis sponge as they assumed it wouldn't work. This was based on limited efficacy of water bath iontophoresis and assumption it would be even less effective via a sponge.

Comparisons of different drugs: Patients felt there was no point in doing this; drugs are mainly the same – some work for some patients, some for others. If one doesn't work or has troublesome side effects, a different one is tried. It would only be worthwhile researching a new drug that had the potential for benefit but with greatly superior side effect profile. Medication most useful when symptoms are at multiple sites – localised hyperhidrosis is easier to control using non-medication interventions (e.g. botox, curettage); more generalised hyperhidrosis (axilla + elsewhere) is harder to treat with something other than medication. Alison Layton

commented that it might be difficult to power a study (for statistical significance) to find differences between medications as they all work quite well. In practice, she will try topical first (fewer adverse effects), then try oral glycopyrrolate, followed by other drugs. It is best to try one, then try a different one, as some work better for some patients than others.

NB Topical glycopyrrolate is missing from the model?

### **Cost-effectiveness of treatment sequences (hyperhidrosis of axilla)**

The sequence Iontophoresis-Botox-Medication-Curettage-Endoscopic Thoracic Sympathectomy was optimal in CE analysis. Patients were happy with this. They felt it was a reasonable sequence. Alison also commented that she might favour botox before medication, particularly in younger patients who don't want the adverse effects of medication. Botox is now more freely available, so happy to try this before medication.

They felt NHS should pay for iontophoresis machines if they work (and are used).

### **QoL tools**

All agreed HidroQol was the best: covers everything important and easy to complete.

DLQI – too general (Alison commented that there is more emphasis on using disease specific tools – hyperhidrosis, rather than general skin disease).

HDSS – very basic (depending on the situation, score can vary between HDSS2 and HDSS3).

Measurement of actual amount of sweating LESS IMPORTANT (should be secondary outcome). Single measurements in time could give wrong impression of level of condition and will not necessarily reflect the patient's overall condition at all. HidroQol should be the primary outcome. Quality of life does generally correlate with sweat rate – but not necessarily at the points in time when measured. There may not be a linear relationship between quantity of sweat produced and quality of life.

**Appendix 4: Collated notes from all hyperhidrosis stakeholder meetings  
(used to develop recommendations for further research)**

Clinical evidence review	Cost-effectiveness analysis	Clinician and patient advisor input	Recommendation for further research
Iontophoresis for HH of the hands has some limited efficacy – a reasonable first option.	Not in CE analysis.	<p>Patients’ experiences reflected trial results – limited effectiveness. Dermatology nurses said good enough that 75-80% don’t come back from further treatment. Dermatologist said this is often seasonal i.e. worse in the hotter weather and agree once controlled some patients will purchase a machine and control their symptoms on an as and when basis. Others move onto other treatment options - 50% of our patients don’t get good enough effect and move on. About 20% purchase a machine and the others just don’t return – not sure of outcomes with these. My concern about saying they don’t come back does not necessarily mean they have been cured of their problem.</p>	None needed.
There is no evidence for iontophoresis (sponge) for HH of axilla.	A trial to establish the efficacy of iontophoresis vs placebo in axilla would be informative for the EVPI analysis of other research questions.	<p>Patients felt it wouldn’t be worth doing further research on iontophoresis sponge as they assumed it wouldn’t work. This was based on limited efficacy of water bath iontophoresis and assumption it would be even less effective via a sponge. Nurses said good enough that 75-80% don’t come back from further treatment. Dermatologist said not sure what we can</p>	No clinical need for a trial, though it would be useful to resolve some methodological uncertainty.

		interpret from not coming back. How many purchased machines – how many had oral anti-cholinergics prescribed?	
Some poor quality evidence that botox has greater efficacy than iontophoresis for HH of the hands A trial of botox versus iontophoresis is HH of the hands is warranted.	Not in CE analysis.	Patients said botox for hands should be studied BUT pain on administration would have to be controlled and patients would need evidence and assurance that botox would not result in long term impairment of hand sensitivity. Dermatologist said different mechanisms to control pain would be helpful – Entonox used successfully and clearly a cost-effective option in the NHS – would be useful to include in a trial.	Trial warranted due to lack of clinical evidence and high unmet need for effective interventions for HH of hand.  Lack of EVPI analysis reflects the lack of evidence – indicating huge uncertainty.  One area that has not been included is iontophoresis using botox – it is more challenging as it is a larger molecule but there are some studies to suggest it may be helpful – this might be something to consider in HH palms and soles?
Topical glycopyrrolate has some limited efficacy in axillary and facial hyperhidrosis.	Not in CE analysis.  Not clear what comprises topical in clinical practice pharmacy might make up some cream. Trials used wipes.	Topical glycopyrrolate (to hand and other areas) didn't have the adverse effects of oral medications (dry eyes/mouth).	Various formulations of glycopyrrolate likely to be very expensive so given low cost of propanthelene bromide not worthwhile.
The trial evidence for oral medications is limited; no clear evidence compared with	The EVPI for medication vs placebo is high – estimates range from £1.4m to £24m.	Clinicians and patients are familiar with these oral medications; moderate efficacy and troublesome side effects. The question of the relative	No clinical value in resolving the quantitative uncertainty – trial of old medications not warranted.

<p>placebo or regarding the relative effectiveness of the various drugs.</p>		<p>effectiveness is not important to patients. Patients think research would be warranted only if a new drug had the potential to be much more effective than the existing ones AND with a much better side effect profile. Many clinicians default to glycopyrrolate despite its high cost believing it to be more effective than other anticholinergics. A trial might be useful to clarify things. However, from a clinical practice perspective easy to try very low cost propantheline first.</p>	
<p>Insufficient evidence for topical botox.</p>	<p>Not analysed.</p>	<p>Not raised as important.</p>	<p>Trial not warranted.</p>
<p>Adequate evidence for efficacy of botox for HH of axilla. Evidence for longer duration (e.g. 1 year) between treatments is lacking.</p>	<p>The EVPI for botox versus placebo is negligible. This reflects the fact that the estimate of the effectiveness of botox vs placebo is fairly precise</p>	<p>Accept that botox is effective and further research is not needed. Patient experience suggests annual treatment is sufficient. The patients agreed that future trials of treatments for hyperhidrosis of the axilla should compare against botox.</p>	<p>Trial not warranted.</p>
<p>No good quality comparison of botox vs curettage.</p>	<p>EVPI suggest such a trial could be cost-effective (though depends on efficacy of iontophoresis for axilla HH).</p>	<p>Patients interested in this comparison. However, the research would have to investigate how permanent the result from curettage is and what is the risk and extent of scarring. Using a form of curettage that minimised scarring but maintained efficacy would be of interest.</p>	<p>Trial <b>NOT</b> warranted yet. Best to await better evidence about which of the treatments aimed at destroying the sweat glands (various types of curettage, laser, microwave, etc.) is best and then compare this with botox. NB important to determine if these</p>

			‘destructive’ therapies offer a permanent cure and offer a good patient outcome in terms of scarring, post-op complications etc.
No evidence comparing botox and oral medication.	EVPI suggest such a trial could be cost-effective (though depends on efficacy of iontophoresis for axilla HH).	Patients only interested in this comparison if the drug was a new one with the potential to be much more effective than the existing ones AND with a much better side effect profile.	Trial not warranted: propathelene bromide so cheap – but still less cost effective than botox. If evidence emerges from ongoing trials of much better efficacy with newer formulations of glycopyrrolate and oxybutynin, a trial vs botox with concurrent cost-effectiveness analysis might be warranted.
Poor quality comparison of laser vs curettage for HH axilla.	Not in CE analysis.	From patients’ perspective research into permanent treatments (like curettage) that reduced risk of scarring without reducing efficacy would be welcomed (e.g. laser, microwave, etc).	Further evidence needs to be generated by developers of new technologies to further demonstrate if any real potential for patient benefit.
Some evidence to favour less radical forms of curettage.	Not in CE analysis.	From patients’ perspective research into permanent treatments (like curettage) that reduced risk of scarring without reducing efficacy would be welcomed (e.g. laser, microwave, etc).	Further evidence needs to be generated by developers of new technologies to further demonstrate if any real potential for patient benefit.
Very limited evidence for microwave and radiofrequency and ultrasound	No evidence to allow these treatments to be modelled – so great uncertainty but	From patients’ perspective research into permanent treatments (like curettage) that reduced risk of scarring without reducing efficacy	Trial may be warranted – further research by sponsors of the technologies may be required to further

technologies in HH.	not yet quantifiable.	would be welcomed (e.g. laser, microwave, etc).	demonstrate if any real potential for patient benefit.
	<p>For axillary HH the most cost-effective sequence is Iontophoresis-Botox-Medication-Curettage-Endoscopic Thoracic Sympathectomy.</p> <p>For axillary HH the lowest estimate of EVPI is £12m.</p>	<p>Patients and clinicians were happy with this. They felt it was a reasonable sequence. They felt NHS should pay for iontophoresis machines if they work (and are used).</p>	<p>NA.</p>

## Appendix 5: Advertisement to recruit patient representatives to the project advisory group for the hepatocellular carcinoma project



UNIVERSITY  
*of York*

**NHS**  
The Leeds  
Teaching Hospitals  
NHS Trust

***Would you like to be involved in research?***

### **Ablative and Non-Invasive Therapies for Hepatocellular Carcinoma study**

#### **Introduction**

The University of York's Centre for Reviews and Dissemination (CRD) is looking for two people with lived experience of liver cancer to join a research project advisory board.

'Lived experience' means anyone who has had, or currently has, hepatocellular carcinoma (HCC), anyone who has cared for, or is currently caring for, someone with HCC.

Hepatocellular carcinoma is a common type of liver cancer. This project will do a systematic review comparing the effectiveness of different ablative and non-invasive therapies for patients with small liver tumours.

A systematic review is a rigorous way of looking at the best available evidence on a particular topic. 'Ablative' and 'non-invasive therapies' are techniques used to destroy tumours without using surgery.

There are many treatments for treating small liver tumours in patients with early stage liver cancer and preserved liver function, including surgery and ablative and non-invasive therapies. However, there are no studies comparing all of these treatments with each other.

#### **What is the study?**

We will identify high quality studies and compare the effectiveness of the different treatments using 'network meta-analysis'. We will assess outcomes that are important to patients, including survival, adverse effects and quality of life.

#### **Who is doing the research study?**

Professor Alison Eastwood, University of York and Dr Ian Rowe, Leeds Teaching Hospitals NHS Trust.

You can read more about Alison's research here:

<https://www.york.ac.uk/crd/staff/alison-eastwood/>

You can read more about Ian's work and research here:

<https://www.leedsth.nhs.uk/a-z-of-services/leeds-liver-unit/meet-the-team/>

### **Who is funding the research?**

The research is funded by the National Institute for Health Research (NIHR) Health Technology Assessment programme (HTA), which funds research looking at the clinical and cost-effectiveness and broader impact of healthcare treatments and tests.

### **What will it involve?**

Being a member of the project advisory group will involve attending two meetings. The first will be to discuss the research protocol and the second will be to discuss the initial findings. A research protocol is a full description of the study, a 'manual' for the team to follow and adhere to, detailing the methods to be used, etc.

We are looking for two people with lived experience of HCC to join the patient collaborator and together to provide advice to the project team based on your lived and personal experience. The other members of the advisory group will include clinical experts from the NHS and researchers from University of York.

### **Time commitment:**

Attend two meetings (lasting 2 hours maximum) one in the coming weeks and one in the summer.

The meetings will be held using 'Zoom' video conferencing software. If you are unfamiliar with Zoom, but happy to use it, we can support you in using the software. Alternatively, if you would prefer to talk on the telephone then we can arrange a separate meeting for this.

### **Criteria:**

We are looking for people over the age of 18 with lived experience of HCC.

### **Will there be reward and recognition (payment) for my time?**

Attendance at each meeting (along with any preparatory reading) should take approximately half a day, for which we will pay £75 for your time.

### **I am interested. Who do I contact?**

If you are interested or if you have any questions, please get in touch with Ros Wade, Research Fellow, University of York. Email:

[ros.wade@york.ac.uk](mailto:ros.wade@york.ac.uk)

## **Appendix 6: Information for patient representatives on the project advisory group for the hepatocellular carcinoma project**

**UNIVERSITY** *of York*  
Centre for Reviews and Dissemination

### **Ablative and non-invasive therapies for early and very early hepatocellular carcinoma: a systematic review and network meta-analysis**

#### **Information for patient representatives on the project advisory group**

Thank you for agreeing to assist us in our study by being a member of the project advisory group. Your input will help us to understand the experience of patients and contribute to the research in this area. This project aims to compare the effectiveness of different treatments for patients with small liver tumours by conducting a systematic review.

A systematic review is a rigorous way of looking at all the best available evidence on a particular topic. By locating, quality assessing and combining the best available research, systematic reviews provide a reliable assessment of what is known and not known.

#### **Background**

Hepatocellular carcinoma (HCC) is the most common form of liver cancer. Patients often have underlying liver disease which can cause physical problems and reduce length of life. Patients, particularly those with advanced HCC, have reduced quality of life and a poor prognosis.

The choice of treatment for HCC depends on a number of factors including where the tumour is located, how big it is, how well the liver is functioning and the general health of the patient. There are several different types of treatment which aim to kill the cancer cells. These include making the tumour extremely hot or cold; injecting the tumour with chemicals; using microwaves or lasers, or blocking blood supply to the tumour (these treatments are described as 'ablative' or 'non-invasive' therapies).

There are no studies comparing the different treatments against each other. Therefore, we will identify all the relevant completed research studies to compare how well these treatments work for HCC patients with small tumours. We will focus on outcomes that matter most to patients, including whether treatments help patients live longer or have better quality of life. We will systematically review the research evidence from randomised controlled trials. We will combine the results of these trials using a network meta-analysis (a statistical technique) to compare how well the treatments work, and if possible, to put them in order of which work best. Statistical methods will be used to test which treatments have reliable evidence, and which need further trials.

Where there are not enough high-quality randomised controlled trials, we will look for other types of study that compare two or more treatments. We will assess whether these are of good enough quality to add useful and reliable evidence about the how well these treatments work.

We will hold a workshop towards the end of the project to share our findings with patients and doctors. We will also use the workshop to find out whether there are particular types of treatment that patients and doctors are interested in which line up with areas where new trials might give stronger and more certain evidence to guide decisions about treatment. The workshop will produce collaborative recommendations on which treatments, comparisons and trial outcomes should become priorities for future research. We will also look at whether it will be possible to undertake economic analysis to see whether the treatments offer good value to the NHS.

We will make sure that we tell relevant audiences about the results of our project. We will produce easy to read summaries and use social media to share the main results of our research.

### **Role of the advisory group**

The role of the project advisory group is to provide advice based on clinical expertise and/or personal experience. You have been asked to participate in our advisory group because we would like to try to understand the experience of patients who have liver cancer. We hope to have 2-3 patient representatives in the group. The clinical experts in the advisory group are Dr Rebecca Goody (Consultant in Clinical Oncology), Dr Jai Patel (Consultant Vascular and Interventional Radiologist), Professor Ajith Siriwardena (Consultant Hepatobiliary and General Surgeon) and Dr Tze Wah (Senior Consultant Radiologist in Diagnostic and Interventional Radiology).

As a member of the advisory group you will be invited to attend two advisory group meetings and the workshop. The purpose of the first meeting is to discuss the draft protocol (a document describing the methods we will follow for undertaking this work), for example by telling us what aspects of treatment you think are important to patients. The second meeting will be held in the summer, when we will ask for comments on the initial research findings from a patient's perspective. The workshop will include additional doctors and patients, alongside the advisory group, where we will discuss the project findings and identify priorities for future research. There are no right or wrong answers to the questions we ask, we are just trying to understand the perspective of patients. If we use technical terms during the meetings, please feel free to ask us to explain their meaning – we want you to feel comfortable working with us and for you to find the process interesting and informative.

The meetings are likely to be held over 'Zoom' videoconferencing software (we will give full instructions on how to use this, if you are not familiar with it). However, if you would prefer to meet separately, or are not available at the time of the advisory group meetings, then we can arrange a separate

meeting either using videoconferencing software or over the telephone. Attendance at each meeting (along with any preparatory reading) should take approximately half a day, for which we will pay you £75.

### **Research team**

Our team consists of researchers with skills in systematic reviews and statistical analysis, a hepatologist who is an expert in HCC and its treatment and a patient collaborator. Details of the members of the research team are listed below:

Mrs Ros Wade, Research Fellow, Centre for Reviews and Dissemination

Mr Gary Raine, Research Fellow, Centre for Reviews and Dissemination

Ms Sahar Sharif-Hurst, Research Fellow, Centre for Reviews and Dissemination

Ms Melissa Harden, Information Specialist, Centre for Reviews and Dissemination

Ms Lindsay Claxton, Health Economist, Centre for Reviews and Dissemination

Professor Sofia Dias, Professor in Health Technology Assessment, Centre for Reviews and Dissemination

Dr Mark Simmonds, Senior Research Fellow, Centre for Reviews and Dissemination

Dr Ian Rowe, Consultant Hepatologist, Leeds Teaching Hospitals NHS Trust

Ms Patricia Thornton, Patient Collaborator

Professor Alison Eastwood, Professor of Research, Centre for Reviews and Dissemination

The project is funded by the National Institute for Health Research (NIHR) Health Technology Assessment (HTA) Programme.

## **Appendix 7: Notes from the hepatocellular carcinoma project first advisory group meeting**

Present: Rebecca Goody, Jai Patel, Tze Wah, Ian Rowe, Trish Thornton (patient), Alison, Sofia, Lindsay, Sahar, Gary, Ros

### **General questions/comments relating to the protocol**

Trish asked for a copy of a paper explaining NMA methods.

*Action: Sahar to send paper and Cochrane videos to Trish.*

### **Inclusion criteria: Interventions**

Ros asked if there are any other relevant interventions not currently listed on page 5 of the protocol.

Tze mentioned electrochemotherapy and histotripsy but mentioned that these are evolving technologies so there is unlikely to be any comparative studies available.

Rebecca suggested searching for wider radiotherapy techniques, rather than just SABR.

Ros asked whether there are any therapies that are not appropriate for specific patients (i.e. do disease/patient characteristics rule out some of the therapies, or it is appropriate to compare all therapies against each other in all populations)?

Ian responded that there are no specific contraindications for one intervention versus a different intervention in early stage HCC patients, unlike for patients with late stage disease.

Tze mentioned that cost may be relevant as equipment costs a lot more for some therapies.

*Action: investigate whether appropriate to add electrochemotherapy, histotripsy and wider radiotherapy techniques to the protocol. If so, Melissa to identify relevant search terms and Ros to add to protocol.*

### **Inclusion criteria: Participants**

Ros asked whether there are any additional clinically-relevant subgroups of interest (for subgroup analysis) and which of the specified subgroups should be prioritised for analysis.

Tze mentioned that the size of the tumour is interesting for ablative techniques.

### **Inclusion criteria: Outcomes**

Ros asked which outcomes are most relevant for patients and clinicians? E.g. which adverse events? This will help us prioritise when drawing conclusions (and also ensure we don't miss any important outcomes).

There was some discussion around the outcome 'time to progression' (not just PFS), which is an important outcome. This is defined differently in different trials, and can include both recurrence in the local area and new tumours developing (due to the underlying liver cirrhosis).

Jai mentioned that outcomes of relevance may differ according to the number of tumours (1-2 tumours vs multiple tumours); treatment is less likely to be curative

for patients with multiple tumours. Therefore, treatment intent is likely to be palliative/life prolonging for patients with multiple tumours, therefore OS is a more relevant outcome, whereas PFS is a more relevant outcome for patients with only 1-2 tumours.

Ian stated that unfortunately most patients will die of liver disease even if the HCC is cured, due to the underlying cirrhosis and the reason for the cirrhosis, therefore, OS is the preferred outcome.

Whilst OS is the primary outcome of interest, PFS is also important as other treatment will be given further down the line (affecting OS). Jai mentioned that a recent TACE publication shows better OS as patients also have systemic therapy so it is difficult to see whether the improvement is from the systemic therapy rather than ablative/intraarterial therapy.

In terms of adverse events, Rebecca stated that liver related toxicity is important and depends on the underlying liver disease (Child-Pugh status). Jai stated that major adverse events are more important than minor adverse events – mild TACE-related post-embolisation syndrome is OK, but major post-embolisation syndrome has a significant impact on quality of life. Underlying liver function affects tolerance to procedures and distribution of HCC tumours may impact on the side effect profile (as a wider area of normal liver is treated). Tze listed the following important adverse events: bleeding (coagulation profile is important), death, pneumothorax, post-ablation syndrome, pain, thermoablative injury (e.g. to the bowel). Jai suggested asking patient groups which outcomes are important, as clinicians' interpretation of important adverse events may differ from a patient's.

*Action: Sahar to amend the data extraction form to include 'time to progression' as well as PFS and space to add the study's definition of 'progression'. Also add a column for recording additional treatments received after the intervention under investigation?*

### **Inclusion criteria: Study location**

Ros asked how applicable the Asian studies are and whether it would be appropriate to pool European and Asian studies.

Tze mentioned the heterogeneity in practice between Asia and Europe, China has a very different way of treating patients.

Jai stated that there is heterogeneity in different areas, not just Europe vs Asia. Practice differs between Leeds and Birmingham, Italy vs. UK, Asia vs. USA. European centres differ too. Aetiology may not give more uniformity, although outcomes may differ by aetiology.

Ian mentioned that Hep B treatment prevents death due to decompensation so studies with patients with primarily Hep B related liver disease will differ from the European population. A lot of patients in Japan are cured from Hepatitis with interferon.

### **Dissemination**

Ros asked whether there are particular groups where we should disseminate our findings.

Ian mentioned the British Association for the Study of the Liver (BASL) – HCC UK have an annual meeting that would be a good forum to share results. The British Liver Trust and Guts UK are relevant patient groups.

Tze mentioned BSIR (British Society of Interventional Radiology) and SIO (Society of Interventional Oncology).

Jai mentioned the CIRSE (Cardiovascular and Interventional Radiological Society of Europe) conference.

Rebecca mentioned the annual SABR (Stereotactic Ablative Radiotherapy) meeting.

### **Workshop attendees**

Ros asked for recommendations for additional clinicians and patients for the workshop in November, to discuss findings and identify key priorities for future research.

Trish stated that the information is very technical and that it might be helpful to have a pre-meeting for patients to explain the terms and methods. Alison said that we will produce pre-workshop information.

Tze has a list of patients that she has previously worked with who might be interested.

Rebecca suggested that clinicians summarise the specific treatment they specialise in for information for the research group and patients. Ian has recently presented an overview of the different treatments and will check whether it is OK to share with us.

*Action: follow up with Tze for a list of patients suitable for the advisory group and workshop. Follow-up with Ian (and the other clinicians) for summaries of the specific treatments.*

### **Additional questions**

Trish asked how early stage HCC is diagnosed. Ian said that patients with cirrhosis are screened using ultrasound, as 70% of HCC is in patients with cirrhosis. There are no early detection methods for those without underlying cirrhosis, so they tend to be diagnosed at a later stage. Therefore, most studies of early HCC will be in patients with underlying cirrhosis.

Alison asked about appropriate dates for the next advisory group meeting – early July is preferred, as school holidays are late July. Early November better for the workshop.

Ros asked the clinicians to let us know of any relevant studies they are aware of (ongoing, published or unpublished).

## **Appendix 8: Notes from the hepatocellular carcinoma project second advisory group meeting**

**Present:** Sumayya Anwer, Dave Clarke, Sofia Dias, Ian Doyle, Alison Eastwood, Rebecca Goody, Robert Hodgson, Richard McCabe, Sahar Sharif-Hurst, Mark Simmonds, Emily South, Trish Thornton, Ros Wade, Tze Wah

**Apologies:** Jai Patel, Ian Rowe, Ajith Siriwardena, Ian Teunion (further to the meeting Jai, Ian and Ian all commented on the meeting notes)

### **Introduction**

Alison outlined the purpose of the meeting – to discuss the interim findings of the systematic review of RCT evidence and prioritise interventions where RCT evidence is lacking, but which are of particular practical interest and warrant targeted searching to identify high quality non-randomised studies. In addition, to prioritise the most relevant patient outcomes.

Ros presented information on the findings of the systematic review of RCTs; 37 RCTs were eligible for inclusion in the review. Most RCTs assessed radiofrequency ablation (RFA), percutaneous ethanol injection (PEI), percutaneous acid injection (PAI), microwave ablation or laser ablation. There were no RCTs available for several ablative/non-invasive therapies.

Sahar presented information on the network of RCT evidence for the outcomes ‘overall survival’, ‘progression-free survival’, ‘overall recurrence’ and ‘local recurrence’. Network meta-analysis will be used to assess and rank interventions by comparative effectiveness for each outcome. Sahar described the comparisons for which hazard ratio/relative risk outcome data are readily available, comparisons for which some assumptions would need to be made to compute the outcome data and those comparisons for which strong assumptions would be required.

Sofia gave an introductory presentation on using observational data alongside RCTs in network meta-analysis. She also explained the threshold method for investigating how much data on a comparison would be needed to change the conclusions drawn based on the network meta-analysis.

### **Discussion**

#### **Important interventions that should be prioritised**

Richard commented that it is difficult to prioritise interventions where further research is required, before seeing the effectiveness results. Alison explained that our initial aim is to identify all the evidence first (including non-randomised comparative studies, where there are gaps in the RCT evidence base), before assessing the clinical effectiveness of the interventions, to ensure we have as comprehensive a network as possible, rather than being led by the results.

Tze asked whether it would be worth including histotripsy (an ultrasound based technology) at this stage, as it is a relatively new technology so we are unlikely to find much non-randomised controlled evidence on this technology. Alison said that we would like to identify the relevant technologies at this stage, then we can look for the evidence and we can inform future research recommendations, where the evidence is lacking. What are the important interventions and what is clinically relevant or relevant from a patient’s perspective?

Tze said that small HCCs are routinely treated with microwave ablation, rather than radiofrequency ablation (which had the most evidence available). There is an evidence gap in terms of RCTs of microwave ablation versus radiofrequency ablation, which is an important gap, which could be filled with non-randomised evidence. All operators in the NHS use microwave, rather than radiofrequency ablation now, so evidence is not keeping up with practice.

Trish stated that from a patient's point of view, she would be most interested in interventions that are the least disruptive to her life (and pain threshold). E.g. not those where multiple appointments or repeat treatments are required or more invasive therapies.

Going through the list of interventions where there was no RCT evidence available, Tze stated that it is unlikely that non-randomised evidence is available for cryoablation, irreversible electroporation, high-intensity focussed ultrasound, electrochemotherapy or histotripsy compared with conventional ablation, but it may be worth checking to confirm that. SABR should be compared against conventional ablation (i.e. microwave ablation or radiofrequency ablation). TACE is usually for patients with multiple lesions or larger lesions. SIRT has recently been commissioned by NHS England but again is more for patients with multiple lesions/larger volume.

Rebecca said that SABR (high dose focussed radiotherapy) was commissioned by NHS England last year and can be given as an alternative to conventional ablation, such as RFA. It also has an important role for patients with other health conditions that mean that they are not suitable for an anaesthetic, SABR does not require an anaesthetic, is non-invasive and outpatient based. SABR and wider radiotherapy techniques are an alternative choice where one of the other treatment types are contraindicated, it is routinely offered in Leeds and becoming more widely available through the UK. As Tze said, TACE and TAE are usually for patients with more widespread disease, so unlikely to be much evidence in very small lesions.

Looking at the matrix of RCT evidence, which are the most important comparisons to focus on? Tze said that if we are focussing on small tumours, microwave versus SABR is an important comparison. Tze confirmed that PEI and PAI are not interventions that should be taken forward, as they are very painful for patients and she does not offer them to patients for this reason, although it has been used in the past and may still be offered in Europe, it is no longer routinely offered in the UK.

Rebecca stated that from a radiotherapy perspective, proton beam therapy is of interest for delivering radiotherapy in a select group of patients, it is offered at proton beam centres in Manchester and London. Along with SABR and standard radiotherapy, radiotherapy is sometimes given in combination with TACE; TACE followed by radiotherapy.

### **Outcomes of interest**

Trish noted that patient satisfaction was only reported in one of the included RCTs. She also stated that length of hospital stay/disruption to life is an important outcome to patients, including time before treatment, as some treatments require multiple scans, etc. Rebecca highlighted that distance patients are required to travel for treatment is also relevant.

Richard considered non-recurrence to be one of the most important outcomes. However, if a procedure is less invasive, requiring a shorter length of hospital stay,

you wouldn't mind having to have the procedure repeated, compared with a procedure that required staying in hospital for 3-4 days – it's a balance between recurrence and disruption to life.

Rebecca highlighted the difference between local disease control with disease in a new area. Local control of the area being treated shows that the treatment has been successful. We know that patients with cirrhosis are at risk of developing other areas of cancer within the liver, but this does not mean that the treatment to the original lesion was not a success. Did the studies report on 'time to next treatment', which can be important for patients' quality of life?

Tze commented that progression-free survival looks at the time after successful treatment to a new tumour developing and needing further treatment, and time to local recurrence; it is important to look at the outcome definitions in the papers, as they mean different things in different studies. In addition, when multiple treatment sessions are planned, whether time to progression is assessed from the time of first treatment or the time of the last treatment.

Alison asked if pain is an important outcome. Trish stated that it would be interesting if it can be compared in any way.

Tze described post-ablation syndrome, which is an immune response with flu-like symptoms that develop around 3-10 days after treatment. However, it is not routinely measured, as most adverse events are measured immediately after treatment. Sofia asked whether non-randomised studies are likely to capture different outcomes, such as longer term outcomes.

Rebecca stated that prospective oncology studies are more likely to capture patient reported outcomes/quality of life, but maybe not the comparative studies. Sofia explained why only comparative studies are relevant for the review, as single arm trials cannot be used to compare different interventions.

Rob outlined the outcomes of relevance for an economic model. The model will require survival outcomes (overall survival, progression-free survival, etc) and economic outcomes, including length of hospital stay. The other major outcome for economic modelling is quality of life. Quality of life data doesn't need to directly come from the comparative studies, there are other sources of utility data that can be used to inform the modelling. Another barrier is length of follow up, if studies only followed patients up for a few years.

Trish asked whether patients are given a choice of interventions and whether they are given information about the risk of complications. Tze said that in terms of ablation techniques, some of the treatments are more recent and outcomes are related to operator experience, therefore, individual institutions measure their own complication rates, so they can be measured against the national standard. The likelihood of complications is explained to patients for the different procedures. Rebecca mentioned that some of the complication rates are also dependent on specific patient characteristics, e.g. the location of the lesion, so it is difficult to give complication rates precisely. In Leeds, if there are a number of treatment options available for a patient, the hepatologist will see the patient in clinic to talk through the potential options, then they will often be referred to Tze or Rebecca or a surgeon, who can give more detail about the specific interventions, to help them make decisions.

## **Conclusions**

### **Summary of interventions to take forward**

Microwave ablation versus radiofrequency ablation

Stereotactic ablative radiotherapy (SABR)

Wider radiotherapy techniques (including proton beam therapy)

Laser ablation (emerging therapy, unlikely to find much data)

Cryoablation (used more in South East Asia than in the West)

Irreversible electroporation (emerging therapy, unlikely to find much data)

High-intensity focussed ultrasound (unlikely to find much data)

Electrochemotherapy (an emerging Italian technology, unlikely to find much data)

Histotripsy (currently being evaluated for CE marking, so unlikely to find comparative data; it is ultrasound based and needleless)

### **Interventions not to be taken forward**

Percutaneous ethanol injection (PEI) (very painful for patients, not routinely used in the UK)

Percutaneous acid injection (PAI) (very painful for patients, not routinely used in the UK)

Transarterial chemoembolization (TACE) (not likely to be used in patients with small tumours)

Transarterial embolization (TAE) (not likely to be used in patients with small tumours)

Selective internal radiation therapy (SIRT) (not likely to be used in patients with small tumours)

## **Appendix 9: Information for patient representatives on the project advisory group for the sudden onset severe headache project**

**UNIVERSITY of York**  
Centre for Reviews and Dissemination



### **Management of sudden onset severe headache presenting to the Emergency Department: a systematic review**

#### **Information for patient representatives on the project advisory group**

Thank you for agreeing to assist us in our study by being a member of the project advisory group. Your input will help us to understand the patient experience of sudden onset severe headache in the Emergency Department and contribute to the research in this area. This project aims to assess the effectiveness and acceptability of care plans for patients who go to hospital Emergency Departments with sudden onset severe headache by conducting a systematic review and holding focus groups with patients.

A systematic review is a rigorous way of looking at all the best available evidence on a particular topic. By locating, quality assessing and combining the best available research, systematic reviews provide a reliable assessment of what is known and not known.

As you will know from personal experience, sudden onset severe headache can be a very painful and worrying condition. Most patients who present to the Emergency Department with a sudden onset severe headache will be diagnosed with migraine or other type of 'primary' headache. However, sudden onset severe headaches can be a sign of a more serious condition, such as subarachnoid haemorrhage (an uncommon type of stroke), so patients who present to Emergency Departments with a sudden onset severe headache undergo tests to ensure that their headache has not been caused by something serious. Headache guidelines recommend that patients have a brain scan, and if the result is normal, they may be offered a lumbar puncture (where a sample of fluid is taken from the spine). However, it isn't always clear which patients need a lumbar puncture after having a normal brain scan.

We will identify studies that have looked at different care plans and tests for patients with sudden onset severe headache. We will assess the accuracy of the tests for identifying subarachnoid haemorrhage and other serious conditions, the side effects of the tests, patient preference and costs. We will gather patients' views on the acceptability of different care plans by holding focus groups with headache patients.

We are currently writing the protocol for the project, which is a document describing the methods we will follow for undertaking this work. When the project is complete we will write a report, describing the research findings

and any recommendations for practice or further research. This will be circulated to relevant healthcare professionals and patient groups.

The role of the project advisory group is to provide advice based on clinical expertise and/or personal experience. You have been asked to participate in our advisory group because we would like to try to understand the experience of patients who go to a hospital Emergency Department with a sudden onset severe headache. We hope to have 2-3 patient representatives in the group. Currently the members are: Dr Alex Danecki (Consultant in Emergency Medicine), Dr Martin Kelsey (Consultant in Emergency Medicine), Dr Husnain Ali (Consultant in Emergency Medicine), Dr Prasad Karadi (Consultant in Acute and General Medicine) and Dr Sayan Datta (Consultant Neurologist).

As a member of the advisory group you will be invited to attend two meetings during the project. The purpose of the first meeting is to discuss the draft protocol, for example by telling us what aspects of care you think are important to patients. The second meeting will be held in autumn 2020, when we will ask you to comment on the research findings, from a patient's perspective. There are no right or wrong answers to the questions we ask, we are just trying to understand the perspective of patients. If there are any technical terms that we use during the meetings, please feel free to ask us to explain their meaning – we want you to feel comfortable working with us and for you to find the process interesting and informative.

The meetings may be held face-to-face or over the telephone, either as part of the main advisory group, or individually, if you prefer. Attendance at each meeting (along with any preparatory reading) should take approximately half a day, for which we will pay you £75. In addition, any travel costs will be refunded.

This research will be undertaken by researchers at the University of York, along with doctors in emergency medicine and neurology at Leeds Teaching Hospitals NHS Trust and a patient collaborator. Details of the members of the research team and their roles in the project are as follows:

Mrs Ros Wade, Research Fellow, Centre for Reviews and Dissemination

Mr Matthew Walton, Research Fellow, Centre for Reviews and Dissemination

Professor Alison Eastwood, Professor of Research, Centre for Reviews and Dissemination

Dr Robert Hodgson, Health Economist, Centre for Reviews and Dissemination

Ms Melissa Harden, Information Specialist, Centre for Reviews and Dissemination

Dr Arabella Scantlebury, Research Fellow, York Trials Unit

Dr Taj Hassan, Consultant in Emergency Medicine, Leeds Teaching Hospitals NHS Trust

Dr James Storey, Consultant Acute Physician, Leeds Teaching Hospitals NHS Trust

Dr Marc Randall, Consultant Neurologist and Stroke Physician, Leeds Teaching Hospitals NHS Trust

Mr John Williams, Patient Collaborator

The project is funded by the National Institute for Health Research (NIHR) Research for Patient Benefit (RfPB) programme.

## **Appendix 10: Project summary and agenda for end-of-project meetings with the project team and clinical advisory group members for the sudden onset severe headache project**

### **Systematic review**

#### **Methods**

18 electronic databases were searched in February 2020 for studies of neurologically intact patients presenting to hospital with non-traumatic sudden onset severe headache (reaching maximum intensity within one hour) with a clinical suspicion of SAH. Eligible studies assessed a care pathway for ruling out SAH, including clinical decision rules and diagnostic tests. Studies were assessed for quality using criteria relevant to the study design. The majority of studies were assessed using the QUADAS-2 tool for diagnostic accuracy studies. Cost-effectiveness studies were assessed using the Drummond checklist. Other study designs were assessed using quality assessment tools specifically developed for the review. Where three or more studies assessed the same intervention and were sufficiently similar, they were pooled using meta-analysis. Other studies were summarised narratively.

#### **Results**

15,750 records were identified. 316 potentially relevant studies were ordered for full paper screening and 51 were eligible for inclusion in the review:

- 37 cohort/before and after studies; 12 had a low risk of bias for all domains, the other 25 were at risk of bias. These studies are described in more detail below, according to which aspect of the care pathway they assessed.
- 4 cost-effectiveness studies; all of which had specific quality issues, reducing the reliability of the results. All 4 studies, undertaken from a US Medicare perspective, modelled different diagnostic strategies (LP, CT angiography, MRI/MRA or no further follow up) for patients presenting with thunderclap headache who had a negative CT result. The results suggest that LP is likely to be the most effective and cost-effective strategy, however, their relevance to UK decision makers is limited.
- 3 systematic reviews of variable quality.
  - A review with a low risk of bias assessed specific headache and patient characteristics, physical examination, CSF analysis, CT and clinical decision rules for SAH; the review was published in 2016, therefore, includes fewer studies assessing diagnostic tests/decision rules than our review. The review found that a history of neck pain and neck stiffness on examination were the individual findings most strongly associated with SAH, that CT within 6 hours was highly accurate and that CSF analysis had lower diagnostic accuracy. They concluded that LP appears to benefit relatively few patients and that clinical decision rules to identify subsets of patients most likely to benefit post-CT LP await external validation.
  - A review with an unclear risk of bias assessed CT within 6 hours of headache onset; not all studies included neurologically intact patients with sudden onset severe headache, therefore, findings may not be generalisable to our population of interest. CT within 6

hours of headache onset was found to be extremely sensitive for ruling out aneurysmal SAH.

- The other review was conducted to derive American College of Emergency Physicians clinical policy and not all included studies met our review inclusion criteria; the review had a high risk of bias. The review concluded that the only risk stratification that reliably identifies the need for neuroimaging is the Ottawa SAH Rule, but that it has poor specificity, that CT performed within 6 hours of symptom onset is sufficient to preclude further diagnostic workup for SAH and that CTA appears to be a reasonable alternative to LP to safely rule out SAH.
- 7 surveys explored clinicians' approach to the investigation of patients with sudden onset severe headache. One UK-based survey of unclear quality reported that ED clinicians had a higher risk tolerance for missed SAH diagnosis than neurospecialists, with neurospecialists more likely to advocate routine LPs compared with ED clinicians. Two poor quality UK-based surveys assessed knowledge of acute headache management amongst emergency and acute medicine clinicians and the need for a guideline; 95% of respondents in one of the surveys indicated that they would find a Trust acute headache guideline useful, whilst only 22% of respondents in the other survey were aware of a local protocol for the investigation of acute headache. A large, good quality survey of ED clinicians from Australia, Canada, the UK and the USA aimed to determine ED practice for investigating acute headache and whether clinicians would consider using a clinical decision rule; responses varied between countries and 96% reported that they would consider using a well-validated clinical decision rule to determine the need for investigations to rule out SAH. A good quality survey of ED clinicians in the USA and Canada assessed knowledge of headache management and adherence to clinical policy; responses varied according to site, academic setting and experience level. One Australian survey of unclear quality interviewed ED clinicians to identify factors that influenced their decisions about diagnostic testing for headache patients after a normal brain CT; patient interaction/preference was at the forefront of the identified factors. A poor quality Australian survey of ED clinicians and trainees assessed ED practice on several aspects of the investigation of acute headache.

***Clinical decision rules***

13 studies assessed Canadian clinical decision rules developed by Perry *et al.*: Rules 1, 2, and 3 and the Ottawa SAH Rule (described below); patients require investigation if one or more findings are present. These rules have also been assessed in studies undertaken in the UK, the USA, Australia, Hong Kong and Taiwan. There are no studies of other clinical decision rules for SAH.

Rule 1	Age $\geq$ 40 years; neck pain or stiffness; witnessed loss of consciousness; onset during exertion
Rule 2	Age $\geq$ 45 years; arrival by ambulance; $\geq$ 1 episode of vomiting; diastolic BP $\geq$ 100 mm Hg
Rule 3	Age 45-55 years; neck pain or stiffness; arrival by ambulance; systolic BP $\geq$ 160 mm Hg
Ottawa SAH Rule	Age $\geq$ 40 years; neck pain or stiffness; witnessed loss of consciousness; onset during exertion; thunderclap

	headache (instantly peaking pain); limited neck flexion on examination
--	--

Rules 1, 2 and 3 had a sensitivity of 90-100% for identifying patients with SAH, but specificity was low (27-43%), resulting in a large number of patients undergoing additional tests. A UK study found that local practice had a sensitivity of 100% and specificity of 66%; therefore, the investigation rate would have increased substantially (from 37% to  $\geq 59\%$ ) with the use of Rules 1, 2, or 3. Another UK study reported that whilst no cases of SAH would have been missed using Rules 1, 2, and 3, nine cases of other significant pathologies would have been missed by employing the clinical decision rules (e.g. intra-parenchymal bleeds, tumours and infarction).

Perry *et al.* refined Rule 1 to develop the Ottawa SAH Rule, which had a sensitivity of 100% (in all but the study from Hong Kong), but a specificity ranging from 8-44% (pooled specificity 24%; 8 studies). A UK study found that current practice had a sensitivity of 100% and a specificity of 59%; the Ottawa SAH Rule would have significantly increased the investigation rate (from 43% to 62%). An Australian study found that use of the Ottawa SAH Rule would double the investigation rate (from 39% to 78%). One study which aimed to validate the Ottawa SAH Rule in Asian Chinese patients demonstrated a much lower sensitivity of 94% (specificity 33%).

#### ***Pathway of CT followed by LP***

The pathway of CT followed by LP was assessed in six studies from Canada, the UK, Spain, Italy and the Netherlands. This pathway had a sensitivity of 100% for detecting SAH, although specificity was quite low in some studies, owing to the high false-positive rate for LP. The pathway also identified other significant pathologies, such as intracerebral haemorrhage, brain tumour, and meningitis.

#### ***CT***

The diagnostic accuracy of CT was assessed in nine studies (although three of the Canadian studies had significant patient overlap so only the largest study was included in the meta-analysis). Four studies (from Canada, the Netherlands and Spain) presented diagnostic accuracy data for CT undertaken within 6 hours of headache onset; pooled sensitivity was 99.2% (95% CI: 93-100) and specificity was 100% (95% CI: 99.9 – 100). Three studies (from Canada, the Netherlands and the UK) assessed CT regardless of time interval; pooled sensitivity was 94% (95% CI: 91-96) and pooled specificity was 100%. Two studies reported diagnostic accuracy data for CT undertaken beyond 6 hours of headache onset; 85.7% and 90%. The prevalence of SAH was much higher in the Dutch study included in the meta-analyses (35-42%) owing to patient recruitment methods, therefore, this study population is unlikely to be representative of patients seen in UK practice. Excluding the Dutch study, the prevalence of SAH in studies of CT undertaken within 6 hours of headache onset was 9.2% to 12.7% and in studies of CT regardless of time interval was 2.7% to 6.2%.

A UK cohort study compared the interpretation of CT scans by Emergency Physicians with neuroradiologists; this study was at a high risk of bias owing to different hardware used to view images between specialties.

## ***LP***

The diagnostic accuracy of LP (CSF analysis using either visual inspection or spectrophotometric assessment) was assessed in 11 studies from Canada, the UK, the USA, Sweden, Spain and the Netherlands. Most studies recruited patients who had a normal CT scan result, therefore, the prevalence of SAH was very low in most studies. Visual inspection for xanthochromia had a pooled sensitivity of 85% (95% CI: 60-95) for detecting SAH and a pooled specificity of 98% (95% CI: 95-99); 3 studies reported sufficient data for pooling (population weighted prevalence of SAH 2%). Spectrophotometric inspection of CSF (UK NEQAS) had a pooled sensitivity of 100% and pooled specificity of 95% (95% CI: 86-98); 3 studies (population weighted prevalence of SAH 0.65%). Two studies reported rates of LP-related complications; in one study 9.5% patients returned to the ED with post-puncture headache (2 of them were admitted for pain control) and one study reported that 5.3% of patients had LP-related complications resulting in a return visit to the ED or hospitalisation.

Two Canadian studies compared visual inspection of CSF versus spectrophotometry and an American study attempted to validate a clinical prediction rule to differentiate between traumatic LP and SAH.

## ***CT angiography***

Two Dutch studies assessed CT angiography after normal CT/LP; no cases of SAH were identified, although 6-19% patients had a vascular abnormality identified, including aneurysm, cerebral venous thrombosis, reversible cerebral vasoconstriction syndrome, cervical dissection and ischemia.

## ***History and examination***

Three studies assessed patient assessment using history and examination. A Canadian study and a UK study investigated the adequacy of patient assessment for SAH and a Dutch study assessed neurologic examination for neck stiffness as a predictor of SAH. Using physicians' clinical suspicion (without the use of a clinical decision rule) resulted in missed cases of SAH. Neurologic examination for neck stiffness was a poor predictor of SAH (sensitivity 67%, specificity 89%). Adequacy of recording of history and complete examination in medical records was poor.

## **Focus groups**

Approval for the qualitative study was obtained from Leeds Teaching Hospitals NHS Trust on 31<sup>st</sup> July 2020 and University of York Health Sciences Research Governance Committee on 3<sup>rd</sup> August 2020. In both cases the study was considered a service evaluation. Due to the issues faced in recruiting patients to the qualitative study (see below), approval was also obtained to collect data through qualitative interviews. This will ensure that patients can be contacted and interviewed as soon as their contact details are received by the qualitative team and will avoid any potential delays associated with having to wait for sufficient numbers of patients to conduct a focus group.

Following advice from the clinical co-applicants, two wards (neurology and acute medicine) were identified and set-up to approach patients to the qualitative study in September. A consultant and trainee(s) were identified for each ward to co-ordinate approaching patients to the qualitative study. Staff were asked to approach any patients meeting the following criteria: neurologically intact patients presenting to

hospital with sudden onset severe headache (peaking within an hour), who have undergone CT to rule out subarachnoid haemorrhage – some patients may also have undergone lumbar puncture. Any patients meeting these criteria were given a participant information sheet and asked if they would be happy to participate in a qualitative study. Patients indicating an interest in the qualitative study were asked to complete a consent to contact form. To facilitate this process a ‘staff manual’ was developed and distributed to all staff who were involved in approaching patients to the qualitative study. The manual outlined: what the study was about, which patients should be approached to the qualitative study, how staff should approach patients and processes for storing and sending consent to contact forms for potential participants to the qualitative team. Staff were also provided with copies of the participant information sheet and consent to contact forms. The qualitative team have been in contact with staff involved in approaching patients on a weekly basis via email and/or telephone.

It was always anticipated that recruitment to the qualitative study would be challenging due to the small number of potentially eligible patients in our target population. However, despite employing a range of strategies to maximise recruitment (listed above), the challenges we have faced reflect those associated with recruiting to qualitative studies and of undertaking primary research during the COVID-19 pandemic more broadly. For this study, changes to the clinical pathway and local restrictions have greatly affected patient recruitment and in particular the ‘footfall’ of patients at Leeds Teaching Hospitals NHS Trust. As a result, despite staff actively trying to approach patients since late September, to date only 5 consent to contact forms have been received by the qualitative team. No patients have agreed to be interviewed.

In light of the problems we have encountered to date and the forthcoming tightening of COVID-19 restrictions nationally, the qualitative team are in contact with the clinical teams on site to discuss whether it will be feasible to continue to approach patients during November.

### **Meeting agenda**

- 1) How would clinicians prefer to see statistical results presented, i.e. what is the most relatable to practice and what are you most used to seeing? Should diagnostic accuracy be presented as false positive/negative rates, or sensitivity/specificity? Is ‘number needed to test’ a useful metric?
- 2) We would like your help to interpret the findings of the review. Do the following conclusions/recommendations appear appropriate?

### ***Conclusions on the effectiveness of care pathways for excluding SAH in patients with sudden onset severe headache:***

- The Ottawa SAH Rule is highly sensitive for identifying patients who require diagnostic testing for SAH, but not very specific (pooled false positive rate: 76%), resulting in increased testing in headache patients. A comparison of the Ottawa SAH Rule with UK practice without a clinical decision rule showed significantly higher rates of testing would be required if the Ottawa SAH Rule were introduced.
- CT within 6 hours of headache onset is highly accurate for identifying SAH, if images are assessed by a neuroradiologist or radiologist who routinely interprets brain CT images (pooled sensitivity 99.18%, pooled specificity 99.95%). Around 1017 patients (95% CI: 112 – 9,807) may

need to undergo additional testing to identify one case of SAH in patients who were classed as negative by <6 hr CT.

- Should there be a caveat that this conclusion is only applicable to patients who are not severely anaemic (one of the false negative results in Perry 2020 was in a patient with sickle cell anaemia) – is it clinically plausible that sensitivity will be lower in such patients?
- Do we need to emphasise the fact that in centres where images are not checked by a neuroradiologist or radiologist who routinely interprets brain images, that sensitivity is likely to be lower? Is review by a neuroradiologist/experienced radiologist for the ‘sign off’ report standard NHS practice, or does this differ between Trusts?
- The figure of >1000 requiring additional testing to identify one case of SAH is heavily influenced by the prevalence of SAH in the study populations. Prevalence was much higher in patients who had CT <6 hours (around 10%) compared with CT at any time (around 5%) – does this difference in prevalence seem clinically plausible (i.e. do SAH patients present earlier, get rushed through to CT quicker)?
- LP (spectrophotometric assessment of CSF) is highly sensitive (pooled sensitivity 100%), but had lower specificity (pooled specificity 95% [95% CI: 86 – 98]) due to ‘traumatic LP’ causing false positives. LP is also associated with adverse events (rates of adverse events requiring revisit to ED or hospitalisation were 5.3-9.5%, where reported). In addition to adverse events relating to the LP procedure, adverse events may occur as a result of additional testing required for patients with positive results (such as contrast-related and radiation exposure-related complications of CT angiogram).
- If CT is not performed within 6 hours of headache onset, then it may be appropriate to undertake additional testing (such as LP) in patients where SAH is still suspected. Pooled sensitivity of CT undertaken at any time since headache onset was 94% (95% CI: 91-96).

***Recommendations for further research:***

- No studies were identified assessing LP on an ambulatory basis (for those patients who require LP after negative CT result). Clinical advice indicated variation in practice regarding whether patients remain in hospital until LP is performed and results are received, or whether LP is done on an ambulatory basis. Therefore, it may be appropriate to undertake a primary study to assess undertaking LP on an ambulatory basis (would it be possible to undertake a retrospective casenote review from different Trusts in the first instance?).
- Would it be appropriate to recommend the investigation of a clinical decision rule that may be more specific/appropriate for a UK NHS setting than the Ottawa SAH Rule? Are there local protocols used in current practice that could be compared for sensitivity and specificity?
- There are no cost-effectiveness studies from a UK perspective. Therefore, it may be helpful to undertake an economic modelling study to investigate the cost-effectiveness of different care pathways for patients presenting to hospital with sudden onset severe headache (e.g. whether to undertake LP after negative CT, undertaking LP on an inpatient vs ambulatory basis).

There is a 100-centre UK-based study planned (SHED) which aims to collect data on 9000 headache patients during 2021 to assess the accuracy of CT within 6 hours and at different time points (at hourly intervals from 6-24 hours). The study is being undertaken by the Royal College of Emergency Medicine Trainee Emergency Research Network (Chief investigator: Professor Dan Horner from Salford Royal NHS Foundation Trust, Key investigator: Dr Tom Roberts, Musgrove Park Hospital, Taunton). The results should be available early in 2022.

- Would it be appropriate to highlight the importance of this UK-based study to provide a definitive conclusion for <6 hr CT accuracy (and requirement for additional testing) and also CT accuracy at different time intervals from headache onset.

3) Are there other questions that have not been answered owing to a lack of research evidence, where recommendations for further research should be made?

4) Are you happy to be contacted to respond to specific clinical questions in the draft report (to be sent separately to those members of the advisory group with the relevant expertise)? We will send the full report to the project team (co-authors) for comment and any advisory group members who would be interested in receiving the full report to provide additional comments (in December).

## **Appendix 11: Notes from end-of-project meetings with the project team and clinical advisory group members for the sudden onset severe headache project**

### **Zoom meeting on 10th November 2020**

Present: Mrs Ros Wade, Research Fellow, Centre for Reviews and Dissemination

Mr Matthew Walton, Research Fellow, Centre for Reviews and Dissemination

Professor Alison Eastwood, Professor of Research, Centre for Reviews and Dissemination

Dr Robert Hodgson, Health Economist, Centre for Reviews and Dissemination

Dr Arabella Scantlebury, Research Fellow, York Trials Unit

Mr John Williams, Patient collaborator

Dr Abu Hassan, Consultant in Emergency Medicine, Leeds Teaching Hospitals NHS Trust

Dr James Storey, Consultant Acute Physician, Leeds Teaching Hospitals NHS Trust

Dr Taj Hassan, Consultant in Emergency Medicine, Leeds Teaching Hospitals NHS Trust

Dr Husnain Ali, Registrar in Emergency Medicine, Leeds Teaching Hospitals NHS Trust

Dr Tony Goddard, Consultant Diagnostic and Interventional Neuroradiologist, Leeds Teaching Hospitals NHS Trust

Dr Prasad Karadi, Consultant in Acute and General Medicine, Calderdale and Huddersfield NHS Foundation Trust

Dr Sayan Datta, Consultant Neurologist, York Teaching Hospital NHS Foundation Trust

### **Microsoft Teams meeting on 11th November 2020**

Present: Mrs Ros Wade, Research Fellow, Centre for Reviews and Dissemination

Dr Sarah Forbes, Associate Medical Director, Leeds Clinical Commissioning Group

#### **1) Clinician preference for presentation of statistical results**

Clinicians are used to seeing results presented in terms of sensitivity and specificity, although it would be helpful if conclusions also mention false positive/false negative rates. The 'number needed to test' figure is also useful and can be used in discussions with patients.

## **2) Interpretation of review results and proposed conclusions/recommendations**

### **The Ottawa SAH Rule**

The review found that the Ottawa SAH Rule is highly sensitive for identifying patients who require diagnostic testing for SAH, but not very specific (pooled false positive rate: 76%), resulting in increased testing in headache patients.

Dr Storey agreed that the Ottawa SAH Rule is too broad, resulting in too many patients being tested. There was a discussion around whether there was a correlation between speed of headache onset and SAH diagnosis and whether using a clinical decision rule for patients whose headache peaked within one hour is too broad. Dr Goddard stated that most SAH patients state that their pain peaks instantly (although recall may be unreliable as time progresses).

It was suggested that we could make a recommendation for further research to assess the specificity of the Ottawa SAH Rule in patients with instantly peaking headache; although Dr Ali commented that such a study would need to be prospective, rather than retrospective, due to potentially insufficient reporting of patient history in medical records.

### **CT within 6 hours of headache onset**

The review found that CT within 6 hours of headache onset is highly accurate for identifying SAH, if images are assessed by a neuroradiologist or radiologist who routinely interprets brain CT images (pooled sensitivity 99.18%, pooled specificity 99.95%). Around 1017 patients (95% CI: 112 – 9,807) may need to undergo additional testing to identify one case of SAH in patients who were classed as negative by <6 hr CT.

Dr Ali highlighted the importance of involving the patient in the decision of whether additional testing is required after a negative CT result, i.e. tell the patient that you are 99% sure that they have not suffered a SAH, then give the patient the choice of whether to proceed to LP (along with information on potential adverse effects of LP and of the implications of a missed SAH). Our conclusion should emphasise patient involvement in decision making.

John stated that, as a patient, he would be reluctant to undergo LP if there was only a 1% chance that the CT result was wrong.

Dr Goddard highlighted the inaccuracy of LP; some samples are inadequate and ambiguous findings can necessitate further investigation in healthy patients. If a patient proceeds to angiography which identifies an aneurysm that has not bled, this leads to difficult clinical decisions.

Dr T Hassan confirmed that he would be comfortable not undertaking LP in patients who have a negative CT result within 6 hours (along with clinical judgement based on history, etc), but not those whose CT was undertaken beyond 6

hours from headache onset (the review found that sensitivity of CT >6 hours after symptom onset was 86-90%).

In a subsequent meeting, Dr Forbes highlighted the implications of the reduced sensitivity of CT beyond 6 hours of headache onset for primary care and emergency medicine, in terms of triaging patients for urgent CT.

Ros asked whether it should be emphasised that for centres where CT images are not checked by a neuroradiologist or radiologist who routinely interprets brain images, that sensitivity is likely to be lower. It was confirmed that this differs between centres (radiologist expertise) therefore non-neurology centres are likely to have lower CT sensitivity. Dr Datta commented that it is difficult to make recommendations that are only appropriate for neurosurgical centres, rather than smaller trusts who do not have neurology input. However, it was agreed that neuroradiology/consultant radiologist sign off is required, in order to rely on the accuracy of CT within 6 hours.

Ros asked whether it was clinically plausible that patients with sickle cell anaemia would have lower CT sensitivity and whether there should be a caveat that conclusions about <6 hour CT sensitivity are not applicable to severely anaemic patients. Dr Goddard said this was not something he had come across and it doesn't seem appropriate to make different recommendations for anaemic patients given their small numbers and important differential diagnoses in sickle cell anaemia.

Ros asked about the higher prevalence of SAH in patients who have CT <6 hours from headache onset, compared to those who have CT beyond 6 hours. Dr Goddard stated that there are a lot of factors that can delay CT in headache patients. Patients do not always present quickly, especially at Christmas time or during the COVID pandemic. Also, patients with a lower volume bleed feel less seriously ill; CT is less sensitive in such patients as a smaller amount of blood is more difficult to detect on CT. Non-patient related factors may also delay CT, e.g. other patients requiring urgent CT.

Dr T Hassan and Dr Goddard discussed the dilemma of late presenting patients (days after headache onset). This could be another recommendation for further research (see below); which diagnostic tests to undertake for late presenting patients with symptoms suggestive of SAH (e.g. CTA, MRI, MRA).

### **Ambulatory LP**

No studies were identified assessing LP on an ambulatory basis (for those patients who require LP after negative CT result). Clinical advice indicated variation in practice regarding whether patients remain in hospital or whether LP is done on an ambulatory basis. Therefore, we plan to recommend further primary research to address this question.

Dr Karadi mentioned that at Calderdale patients are offered LP the next day, on an ambulatory basis, where LP is required for reassurance. There was discussion around whether delaying LP until the next day (e.g. if the patient presents during the night) would result in a change in urgent patient management, or whether the delay is acceptable from a neurology perspective.

Dr T Hassan commented that patients who are in severe pain need to be admitted for pain control, but that it is better to LP patients during the day as results are only processed during the daytime anyway. Dr Storey stated that as CT-ve patients who

have a positive LP result would need to have further tests (e.g. angiography) anyway, undertaking LP on an ambulatory basis would be unlikely to change the immediate treatment plan.

### **3) Other recommendations for practice or further research**

Dr T Hassan and Dr Goddard discussed the difficulties associated with diagnosing SAH in patients presenting  $\geq 7$  days after symptom onset, and the lack of guidance & consistency with how these patients are assessed. CT angiography may be more appropriate as xanthochromia may no longer be present, CTA will also pick up other pathologies e.g. RCVS. Although MRI/MRA may be a better option in these patients, as blood products are still visible long after a bleed. Patients presenting late to the neurology department are much more likely to receive MRI/MRA and thus receive diagnosis than those who present to the ED. Guidance needs to be produced for delayed presentation sudden onset severe headache patients based on new primary research.

## **Appendix 12: Patient and Public Involvement section of the sudden onset severe headache RfPB report (following GRIPP2-SF guidance)**

**Aim:** The aim of PPI in this project was to ensure that the patient's perspective was captured at all stages, from protocol development (including deciding which outcomes should be assessed by the systematic review) through to interpreting the results of the review and drawing conclusions.

**Methods:** A patient collaborator with experience of presenting to the ED with sudden onset severe headache was involved throughout the project. Three additional patients were identified by our clinical co-investigators and recruited to the advisory group; all three patient advisors were available at the beginning of the project to advise on the protocol and two of the patient advisors (along with the patient collaborator) were available at the end of the project to help interpret the results. Meetings with the patient collaborator and patient advisory group members were undertaken face-to-face or via telephone or Zoom videoconferencing software, according to patient preference and COVID-19 restrictions in place at the time of the meetings.

**Results:** The PPI aspect of the project added context to the review findings and highlighted preferences of our patient advisors regarding the assessment of sudden onset severe headache; this also informed our recommendation for further primary research on the setting for undertaking LP, when required (inpatient versus ambulatory care).

**Discussion and conclusions:** The input from patients about which outcomes were the most important to them was very informative when developing the systematic review protocol. In addition, the patients' help interpreting the findings of the review and the consistency of patients' preference for LP on an ambulatory basis reinforced the importance of further primary research on this specific question, for which no research evidence was identified. The initial meetings with patients were also informative to help the researchers understand the experience of patients attending the ED with sudden onset severe headache, their concerns and preferences. The characteristics and comorbidities of the patients involved in this project were varied, however, their concerns and preferences were generally consistent.

**Reflective/critical perspective:** PPI was an important aspect of this project, enabling researchers to understand the care pathway for the assessment of sudden onset severe headache from a patient's perspective. Whilst it was difficult to present complex review findings to patient advisors who did not have a background in health care or research, requiring additional time and effort to prepare for patient meetings, the feedback from patients was that the information was presented clearly and the patients enjoyed being involved in the project.

### **Appendix 13: Patient and clinician engagement section of the sudden onset severe headache full project report**

The project team included four clinicians with expertise in emergency medicine, acute medicine, neurology, stroke and headache, and a patient collaborator with experience of presenting to the ED with sudden onset severe headache. Three additional patients who presented to the ED at Leeds Teaching Hospitals NHS Trust with sudden onset severe headache and additional clinicians with expertise in emergency medicine, acute medicine, neurology, neuroradiology and an NHS commissioner were recruited to our advisory group (advisory group members are listed on page 2 of this report).

The patients' and clinicians' perspectives were collected at various points through the project including at team and advisory group meetings and during protocol development. The patients' and clinicians' perspectives were used to help with the interpretation of the results of the systematic review.

Discussions at team meetings highlighted a lack of consistency regarding inpatient versus ambulatory LP; practice varied between (a) undertaking LP on an ambulatory basis, (b) undertaking LP while the patient is still in hospital, but then discharging the patient to the ambulatory care unit while the result is awaited (which can take 2-3 days at a district general hospital) or (c) keeping the patient in hospital until the LP has been undertaken and the result is received.

In November 2020, meetings were held with members of the project team and advisory group to discuss the findings of the project, draw conclusions and make recommendations for further research. Due to COVID-19 restrictions, meetings had to be held via Zoom, Microsoft Teams or telephone, rather than face to face.

Clinical and patient members of the project team and advisory group were unsurprised by the findings relating to the diagnostic accuracy of CT, LP and the Ottawa SAH Rule in neurologically intact adults presenting with non-traumatic sudden onset severe headache (peaking within one hour). They highlighted the importance of involving the patient in the decision of whether additional testing is required after a negative CT result; communicating the level of certainty in the diagnostic test result and possible adverse effects of subsequent diagnostic tests to aid the decision-making process.

Clinicians discussed the variation in practice regarding inpatient versus ambulatory LP, when LP is required for reassurance; two patient advisors and the patient collaborator expressed a preference for ambulatory LP. Owing to the lack of studies assessing the setting for LP, it was felt that further primary research may be useful to address this question.

The difficulties associated with diagnosing SAH in patients who present several days after headache onset was also discussed; there is a lack of guidance and consistency in how these patients are assessed. It was concluded that further primary research would be informative in order to develop guidance for this patient subgroup.

## Included papers

Paper 1: Wade R, Paton F, Rice S, Stansby G, Millner P, Flavell H, Fox D, Woolacott N. Thigh length versus knee length antiembolism stockings for the prevention of deep vein thrombosis in postoperative surgical patients; a systematic review and network meta-analysis. *BMJ Open*. 2016;6:e009456. doi:10.1136/bmjopen-2015-009456.

Paper 2: Wade R, Paton F, Woolacott N. Systematic review of patient preference and adherence to the correct use of graduated compression stockings to prevent deep vein thrombosis in surgical patients. *Journal of Advanced Nursing*. 2017;73(2):336-48. doi:10.1111/jan.13148.

Paper 3: Wade R, Llewellyn A, Jones-Diette J, Wright K, Rice S, Layton AM, Levell NJ, Stansby G, Craig D, Woolacott N. Interventional management of hyperhidrosis in secondary care: a systematic review. *British Journal of Dermatology*. 2018;179(3):599-608. doi:10.1111/bjd.16558.

Paper 4: Wade R, Jones-Diette J, Wright K, Layton AM, Woolacott N. Hyperhidrosis quality of life measures: review and patient perspective. *Journal of Dermatological Treatment*. 2019;30:303-8. doi:10.1080/09546634.2018.1506080.

Paper 5: Wade R, South E, Anwer S, Sharif-Hurst S, Harden M, Fulbright H, Hodgson R, Dias S, Simmonds M, Rowe I, Thornton P, Eastwood A. Ablative and non-surgical therapies for early and very early hepatocellular carcinoma: a systematic review and network meta-analysis. *Health Technology Assessment*. 2023;27(29). doi:10.3310/GK5221.

Paper 6: Walton M, Hodgson R, Eastwood A, Harden M, Storey J, Hassan T, Randall M, Hassan A, Williams J, Wade R. Management of patients presenting to the emergency department with sudden onset severe headache: systematic review of diagnostic accuracy studies. *Emergency Medicine Journal*. 2022;39(11):818-25. doi: 10.1136/emmermed-2021-211900.

# BMJ Open Thigh length versus knee length antiembolism stockings for the prevention of deep vein thrombosis in postoperative surgical patients; a systematic review and network meta-analysis

Ros Wade,<sup>1</sup> Fiona Paton,<sup>1</sup> Stephen Rice,<sup>1</sup> Gerard Stansby,<sup>2</sup> Peter Millner,<sup>3</sup> Hayley Flavell,<sup>4</sup> Dave Fox,<sup>1</sup> Nerys Woolcott<sup>1</sup>

**To cite:** Wade R, Paton F, Rice S, *et al*. Thigh length versus knee length antiembolism stockings for the prevention of deep vein thrombosis in postoperative surgical patients; a systematic review and network meta-analysis. *BMJ Open* 2016;**6**:e009456. doi:10.1136/bmjopen-2015-009456

► Prepublication history and additional material is available. To view please visit the journal (<http://dx.doi.org/10.1136/bmjopen-2015-009456>).

Received 22 July 2015  
Revised 22 December 2015  
Accepted 31 December 2015



CrossMark

For numbered affiliations see end of article.

#### Correspondence to

Ros Wade;  
[ros.wade@york.ac.uk](mailto:ros.wade@york.ac.uk)

#### ABSTRACT

**Objectives:** To assess the clinical effectiveness of thigh length versus knee length antiembolism stockings for the prevention of deep vein thrombosis (DVT) in surgical patients.

**Design:** Systematic review and meta-analysis using direct methods and network meta-analysis.

**Methods:** Previous systematic reviews and electronic databases were searched to February 2014 for randomised controlled trials (RCTs) of thigh length or knee length antiembolism stockings in surgical patients. Study quality was assessed using the Cochrane Risk of Bias Tool. The primary outcome was incidence of DVT. Analysis of the DVT data was performed using ORs along with 95% CIs. The I<sup>2</sup> statistic was used to quantify statistical heterogeneity.

**Results:** 23 RCTs were included; there was substantial variation between the trials and many were poorly reported with an unclear risk of bias. Five RCTs directly comparing thigh length versus knee length stockings were pooled and the summary estimate of effect favouring thigh length stockings was not statistically significant (OR 1.48, 95% CI 0.80 to 2.73). 13 RCTs were included in the network meta-analysis; thigh length stockings with pharmacological prophylaxis were more effective than knee length stockings with pharmacological prophylaxis, but again results were not statistically significant (OR 1.76, 95% credible intervals 0.82 to 3.53).

**Conclusions:** Thigh length stockings may be more effective than knee length stockings, but results did not reach statistical significance and the evidence base is weak. Further research to confirm this finding is unlikely to be worthwhile. While thigh length stockings appear to have superior efficacy, practical issues such as patient acceptability may prevent their wide use in clinical practice.

**Systematic review registration number:** CRD42014007202.

#### Strengths and limitations of this study

- This systematic review used all the available randomised evidence on thigh length or knee length antiembolism stockings to indirectly compare the two stocking lengths.
- Many trials were old and poorly reported and there was substantial variation in terms of patient characteristics and interventions used.
- Standard meta-analysis and network meta-analysis were undertaken in order to compare all relevant treatments with one another.
- The results of the network meta-analysis were consistent with the direct meta-analysis, although there was significant statistical heterogeneity in the models.
- The uncertain quality of many of the included trials reduces the reliability of the results of the review.

#### INTRODUCTION

Deep vein thrombosis (DVT) is a condition in which a blood clot forms in one of the deep veins of the body, usually in the leg. An emboli is formed if the blood clot or part of the blood clot detaches and travels through the venous system. If the clot lodges in the lung, this is termed a pulmonary embolism (PE) and this may be fatal. DVT and PE are collectively known as venous thromboembolism (VTE). The House of Commons Health Committee reported in 2005 that an estimated 25 000 people in the UK die each year from potentially preventable hospital-acquired VTE.<sup>1</sup>

Surgical patients are at an increased risk of developing DVT, due to stasis in venous blood flow and increased coagulability of the

blood, caused by factors such as immobilisation, decreased fluid intake and blood or body fluid loss. It has been estimated that between 45 and 51% of patients undergoing orthopaedic surgery develop DVT, if not provided with adequate prophylaxis.<sup>1</sup> Routine prophylaxis reduces morbidity, mortality and health service costs in patients at risk.<sup>2</sup> Prophylaxis can be pharmacological (such as low-molecular-weight heparin (LMWH)) and/or mechanical (such as antiembolism stockings (also known as graduated compression stockings)).

Antiembolism stockings are available as thigh length or knee length stockings. They exert graded pressure at a decreasing gradient from the ankle towards the thigh or knee, which increases blood flow velocity and promotes venous return. Patients have reported that both thigh length and knee length stockings are difficult to use, but fewer patients reported discomfort with knee length stockings and patients are more likely to wear knee length stockings correctly.<sup>3–5</sup>

The National Institute for Health and Care Excellence (NICE) guideline 'Venous thromboembolism: reducing the risk' (CG92) states that the length of stockings is a controversial issue and there is no clear randomised evidence that one length is more effective than another.<sup>6</sup> The Scottish Intercollegiate Guidelines Network (SIGN) guideline on the prevention and management of VTE (SIGN guideline 122) states that studies comparing above-knee with below-knee stockings have been too small to determine whether or not they are equally effective.<sup>2</sup>

This systematic review aims to address this question more definitively by utilising all the available randomised evidence on thigh length or knee length stockings, rather than just trials that directly compare the two stocking lengths: using both standard meta-analysis and network meta-analysis. Network meta-analysis enables a comparison of all relevant treatments with one another. This review was undertaken as part of a larger research project to establish the expected value (cost-effectiveness) of undertaking additional research comparing the relative effectiveness of the two different lengths of stocking, in addition to standard pharmacoprophylaxis.<sup>7</sup>

## METHODS

We conducted a systematic review to assess the clinical effectiveness of thigh length versus knee length antiembolism stockings for the prevention of DVT in surgical patients. Owing to the anticipated paucity of research evidence directly comparing thigh length stockings with knee length stockings, we also sought studies comparing thigh length stockings with a control treatment and studies comparing knee length stockings with a control treatment.

Clinical advice was provided by an advisory group which included a vascular surgeon, an orthopaedic surgeon and an anticoagulant and thrombosis

consultant nurse. A patient representative also provided information on her experiences of using antiembolism stockings after two different types of surgery.

The research protocol was registered on the international prospective register of systematic reviews (PROSPERO registration number: CRD42014007202).

## Search strategy

Eleven guideline and systematic review databases (including the Cochrane Database of Systematic Reviews, Database of Abstracts of Reviews of Effects, PROSPERO, Health Technology Assessment Database and National Guidelines Clearinghouse) were searched up to August 2013 for reviews of antiembolism stockings. The included and excluded studies listed by relevant systematic reviews were screened for relevant primary studies. To update the searches undertaken in the relevant reviews, systematic searches for RCTs published since January 2010 were undertaken in February 2014. Six electronic sources were searched (MEDLINE, MEDLINE In-Process, EMBASE, CINAHL, AMED and CENTRAL) as well as two grey literature databases (ClinicalTrials.gov and Current Controlled Trials). No language restrictions were applied. In addition, clinical advisors were consulted for additional potentially relevant studies and reference lists of all included studies were manually searched. Records were inserted into an EndNote library.

The search strategy developed for Ovid MEDLINE is presented below.

Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R), 1946 to Present. Searched on 19 February 2014. Date limited to 2010 onwards. Search strategy:

1. exp "embolism and thrombosis"/ (172610)
2. (thrombos\$ or thrombus\$ or thrombotic or thrombolic\$ or thromboemboli\$ or thromboprophyla\$ or embol\$).ti,ab. (232741)
3. (DVT\$ or PE or PTS).ti,ab. (34899)
4. 1 or 2 or 3 (317779)
5. Stockings, Compression/ or Compression Bandages/ (1165)
6. (stocking\$ or hose or hosiery or tights or sock\$ or TEDS).ti,ab. (10451)
7. (compression adj3 bandage\$).ti,ab. (486)
8. 5 or 6 or 7 (11541)
9. 4 and 8 (1418)
10. randomized controlled trial.pt. (362662)
11. controlled clinical trial.pt. (87530)
12. randomized.ab. (282970)
13. placebo.ab. (149727)
14. drug therapy.fs. (1661607)
15. randomly.ab. (205717)
16. trial.ab. (291784)
17. groups.ab. (1315795)
18. 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 (3250729)
19. 9 and 18 (518)
20. limit 19 to yr="2010 -Current" (141).

## Study selection

RCTs assessing thigh length or knee length antiembolism stockings (with or without standard pharmacological prophylaxis) in surgical patients were eligible for inclusion; the length of stocking had to be clearly stated. The primary outcome was incidence of DVT; DVT data were included only if definitively diagnosed using radioiodine (<sup>125</sup>I) fibrinogen uptake, venography, Doppler ultrasound or MRI. Studies reporting complications and consequences associated with DVT (such as the incidence of PE, incidence of post-thrombotic syndrome and mortality) or adverse effects related to the use of antiembolism stockings were also included.

Studies identified by the searches were independently assessed for inclusion by two reviewers using the prespecified inclusion criteria stated above. Disagreements were resolved through discussion and, where necessary, by consultation with a third reviewer.

## Data extraction

Data extraction was conducted by one reviewer using a piloted and standardised data extraction form in Eppi-Reviewer 4.0 and independently checked by a second reviewer. Discrepancies were resolved by discussion, with involvement of a third reviewer when necessary. In cases where the same study was reported in multiple publications, the most up to date or comprehensive publication was used for data extraction. Data were extracted on study details (eg, author, year, location of study), patient characteristics (eg, age, gender, type of surgery, baseline risk factors for VTE), details of the intervention (eg, type of stocking, duration of use, co-interventions including pharmacological thromboprophylaxis), and reported outcomes (eg, method of assessment and results).

## Quality assessment

The quality of the individual trials was assessed by one reviewer, and independently checked by a second reviewer; disagreements were resolved by consensus and if necessary a third reviewer was consulted. The quality of included trials was assessed using the Cochrane Risk of Bias Tool, which assesses methods of randomisation and allocation concealment, blinding, completeness of outcome data and selective outcome reporting.<sup>8</sup> Similarity of treatment groups at baseline was also assessed. Each trial was given an overall risk of bias judgement; trials that had a low risk of bias for all key domains were judged to have a low overall risk of bias, trials that had a high risk of bias for one or more key domains were judged to have a high overall risk of bias, and trials that had an unclear risk of bias for one or more key domains were judged to have an unclear overall risk of bias.

## Synthesis

Analysis of the DVT data was performed using ORs along with 95% CIs. Owing to the clinical and methodological variation between trials a random effects model

was used to pool data. The  $I^2$  statistic was used to quantify statistical heterogeneity. The statistical package used for analysis was RevMan V.5.2.

A network meta-analysis (NMA) was performed to investigate whether the utilisation of indirect evidence would increase the precision of the relative effect estimate for thigh length versus knee length stockings. It also provides an estimate of the relative effect of all treatments relative to one another. A high level of inconsistency between the direct and indirect evidence suggests clinical or methodological heterogeneity, which increases the uncertainty in the effect estimates. Although several outcomes were investigated in the review, there was only sufficient evidence to perform an NMA for the outcome DVT. To create the network, interventions that were considered sufficiently similar relative to the interventions of interest were lumped together: the effectiveness of LMWH, low dose heparin and fondaparinux were assumed to be the same, and these were therefore lumped together in the network and were referred to collectively as 'heparin'. Based on the advice of the clinical advisors, it was assumed that there was no stocking-heparin interaction in the base case analysis, that is, the effect of thigh length stockings compared to knee length stockings is the same as thigh length stockings plus concomitant heparin compared to knee length stockings plus concomitant heparin. This assumption was tested in a sensitivity analysis. A random effects analysis was used and credible intervals (CrI) represent the uncertainty around the average treatment effect across trials. The only potential effect modifier for which there was evidence across the trials and a relevant network, was whether or not patients had undergone orthopaedic surgery, which carries a high risk of DVT. Therefore, a subgroup analysis was conducted to compare the effectiveness of antiembolism stockings in orthopaedic surgery patients versus other surgery patients. The model, written in WinBUGS, was based on code presented in the NICE Technical Support Document 2.<sup>9</sup>

Data on the incidence of PE, mortality and adverse events related to the use of antiembolism stockings were tabulated and synthesised narratively.

## RESULTS

During protocol development, scoping searches identified two particularly relevant Cochrane reviews.<sup>10 11</sup> Therefore, many relevant trials were identified from the included and excluded studies lists of these reviews (among others), prior to running the update searches for primary studies.

The electronic search of the relevant systematic review and guideline databases identified 307 records, of which 12 appeared to be systematic reviews of antiembolism stockings in postoperative surgical patients (including the two reviews identified during the protocol development stage). These reviews were obtained so that their lists of

included and excluded studies could be systematically searched for potentially relevant primary studies. A total of 137 records were added to the EndNote library from the included and excluded studies lists of the 12 relevant systematic reviews (after removal of duplicates). The update searches of electronic databases (from 2010 to February 2014) identified an additional 330 records, which were also added to the EndNote library.

The full papers of 68 potentially relevant primary studies were screened for inclusion in the review. Twenty-three RCTs met the inclusion criteria and were included in the systematic review (see online supplementary tables S1–5 for study details).<sup>12–34</sup> Figure 1 presents the flow of studies through the study selection process.

Of these 23 RCTs, 21 reported data for the outcome DVT. However, one trial did not report sufficient data to be included in the meta-analysis or NMA, as total numbers of patients in treatment groups were not reported.<sup>34</sup> Figure 2 shows the network of 20 trials that presented adequate data on DVT.

Seven trials did not add to the network of evidence comparing thigh length with knee length antiembolism stockings: these trials compared thigh length or knee length stockings with a different intervention, such as pneumatic compression or dextran.<sup>14 16 19 24 27 28 31</sup> Both thigh length stockings and knee length stockings needed to be compared with a common comparator to be able to inform the relative effectiveness of the two different stocking lengths. Therefore, 13 RCTs contained data that directly or indirectly informed the relative effectiveness of thigh length versus knee length stockings and were included in the standard meta-analysis or NMA or both.<sup>12 13 15 17 18 20–23 25 29 32 33</sup>

There was substantial variation between the 23 included trials in terms of the patient characteristics, suggesting that the participants had a different baseline risk for DVT. There was also variation in the interventions used in the RCTs; in some trials a stocking was only worn on one leg, rather than both legs, and the duration of use varied between trials. Concomitant pharmacological prophylaxis also varied between trials.

Generally the trial methods were poorly reported, with a high proportion of assessments for each quality domain having to be recorded as unclear. Overall 3 RCTs can be considered to have a low risk of bias,<sup>14 23 25</sup> 5 have a high risk of bias<sup>18 22 27 32 33</sup> and for 15 RCTs the reporting was inadequate to judge the risk of bias.<sup>12 13 15–17 19–21 24 26 28–31 34</sup>

Many of the included RCTs dated back to the 1970s<sup>13 19</sup> and 1980s,<sup>14 16–18 20 22 24 26–31 33</sup> therefore, their results may not be generalisable to current practice; surgical practice has changed over time with less invasive surgical procedures, shorter duration of hospitalisation and earlier mobilisation after surgery.

## DVT results

Twenty RCTs reported rates of DVT and provided sufficient data to be included in meta-analyses. Where

reported, the majority of DVTs were asymptomatic, the clinical consequences of which are unknown.

### Thigh length stockings (with or without pharmacological prophylaxis) versus knee length stockings (with or without pharmacological prophylaxis)

Two RCTs<sup>12 25</sup> directly compared thigh length versus knee length stockings, plus pharmacological prophylaxis, reflecting current practice for the treatment of patients at high risk of DVT; the results were inconsistent in terms of the direction of effect. The reasons for the inconsistent findings between the two trials were unclear and may be due to chance.

Four additional RCTs that compared thigh length versus knee length stockings were identified, but these trials did not include additional pharmacological prophylaxis.<sup>29 32–34</sup> Unfortunately, the trial by Ayhan (2013) was reported only as an abstract and did not provide details on the number of patients in each treatment group; therefore this trial was excluded from meta-analyses.<sup>34</sup>

The five available RCTs comparing thigh length versus knee length stockings with or without additional pharmacological prophylaxis were combined using meta-analysis (figure 3); the summary estimate of effect indicated a trend favouring thigh length stockings, but the findings were not statistically significant (knee vs thigh OR 1.48, 95% CI 0.80 to 2.73,  $p=0.21$ ;  $I^2=33\%$ ).

There was some inconsistency in the direction of effect for trials assessing patients in similar surgical groups. Cohen *et al*<sup>25</sup> and Hui *et al*<sup>32</sup> included orthopaedic patients, and Porteous *et al*<sup>29</sup> and Williams *et al*<sup>33</sup> included patients undergoing abdominal surgery. The reasons for the inconsistency were unclear and may be due to chance.

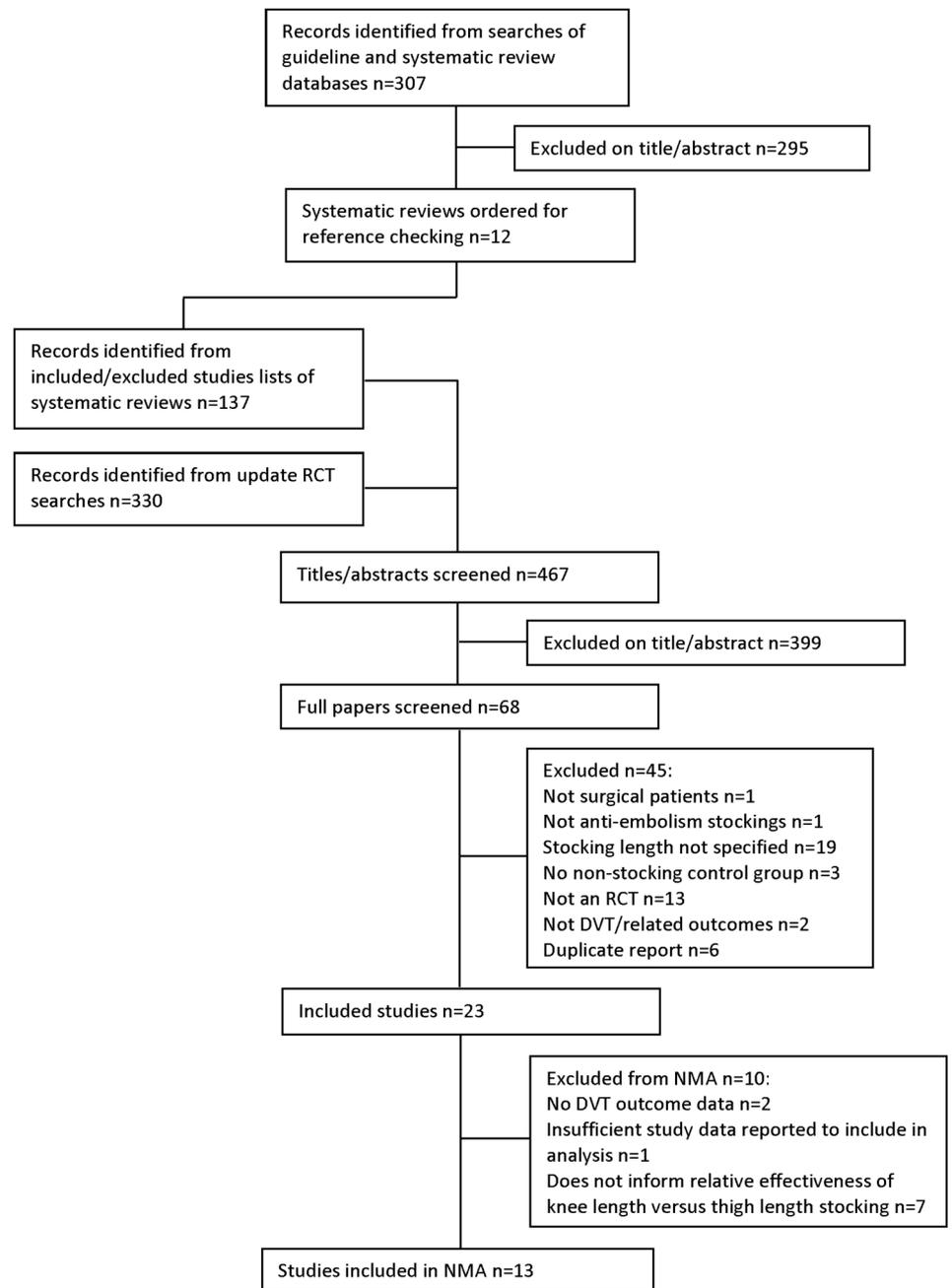
The other 15 RCTs that reported DVT results compared either thigh length or knee length antiembolism stockings with no stocking or with another method of thromboprophylaxis, therefore their results do not directly inform the comparison of thigh versus knee length stockings.

### Network meta-analysis

Thirteen trials contained data that directly or indirectly informed the relative effectiveness of thigh length versus knee length stockings with or without pharmacological prophylaxis for the prevention of DVT and were included in the NMA. Table 1 presents the direct comparisons included in the NMA, and the number of studies reporting that direct comparison. The number of direct comparisons, 19, is greater than the number of studies in the NMA because three three-armed trials were included in the analysis.

The results of the NMA are the estimates of the average effects across a heterogeneous set of trials. The credible intervals (CrI) presented represent the uncertainty around that average. There was significant statistical heterogeneity in the models and inconsistency

**Figure 1** Flow diagram of the study selection process. DVT, deep vein thrombosis; NMA, network meta-analysis; RCT, randomised controlled trial.



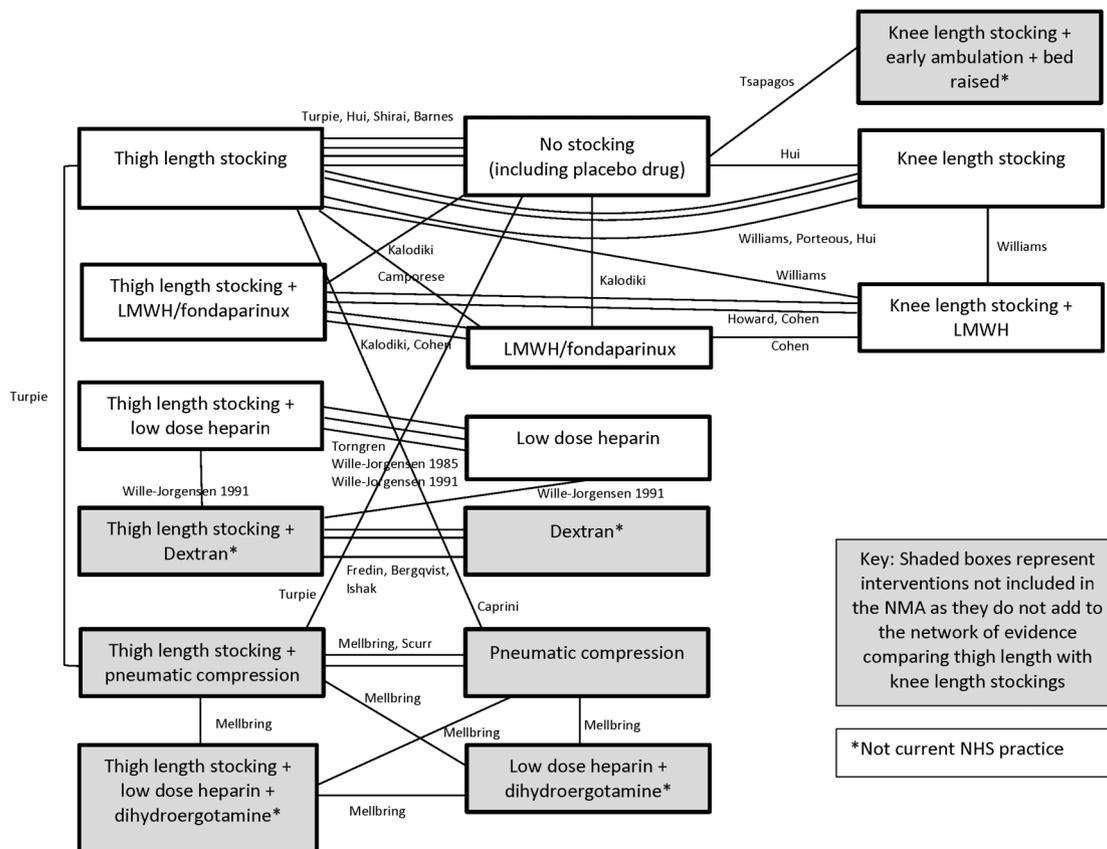
indicating that there may be underlying unknown clinical and methodological heterogeneity across the trials.

The results of the base case analysis found that heparin was statistically significantly more effective than no treatment (median OR 0.26, 95% CrI 0.09 to 0.87,  $p=0.03$ ), thigh length stockings with heparin were statistically significantly more effective than heparin alone (median OR 0.38, 95% CrI 0.21 to 0.63,  $p=0.00$ ) and knee length stockings with heparin were more effective than heparin alone, although this result was not statistically significant (median OR 0.68, 95% CrI 0.27 to 1.38,  $p=0.28$ ).

In the base case analysis, thigh length stockings with pharmacological prophylaxis were more effective than knee length stockings with pharmacological prophylaxis

(knee vs thigh OR 1.76, 95% CrI 0.82 to 3.53,  $p=0.12$ ), but this result was not statistically significant. The indirect estimate favours thigh length stockings slightly more than the direct estimate of 1.48 (95% CI 0.80 to 2.73,  $p=0.21$ ) from the direct meta-analysis presented above, but there is also greater uncertainty in the estimate. The NMA did not increase the precision of the relative effect estimate for thigh length versus knee length stockings because of the uncertainty associated with the inconsistency between direct and indirect estimates of effect. The full table of results in the base case are presented in online supplementary table S6.

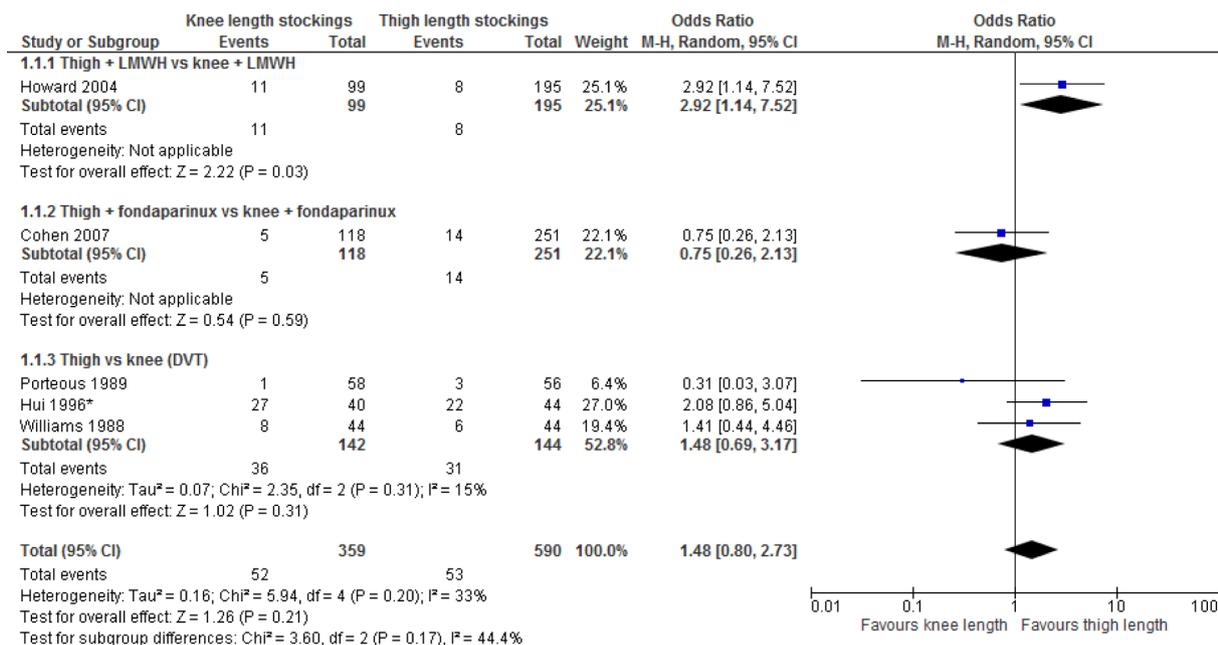
The effectiveness of each treatment is represented by the absolute risk of DVT in table 2. The baseline risk of DVT (both symptomatic and asymptomatic) for



**Figure 2** Network of trials presenting data on DVT. DVT, deep vein thrombosis; LMWH, low-molecular-weight heparin; NHS, National Health Service; NMA, network meta-analysis.

moderate risk general surgical patients taking heparin was estimated to be 9.88%, estimated using the American College of Chest Physicians Guidelines for the prevention of VTE in non-orthopaedic surgical

patients.<sup>35</sup> Using this baseline risk estimate, the absolute risks of DVT for patients using the different treatments are presented below in table 2. The combination of thigh length stockings with pharmacological prophylaxis



\*Number of events calculated from % of DVTs reported and it is unclear whether any were bilateral and therefore whether double counting of patients has occurred.

**Figure 3** Rates of DVT (or VTE) comparing thigh length stockings (with or without pharmacological prophylaxis) versus knee length stockings (with or without pharmacological prophylaxis). DVT, deep vein thrombosis; VTE, venous thromboembolism.

**Table 1** The direct comparisons included in the network meta-analysis, and the number of studies reporting that direct comparison

Treatment	Knee length stocking	No prophylaxis	Knee length stocking plus heparin	Heparin
Thigh length stocking	3	4	1	1
Knee length stocking		1	1	–
Thigh length stocking plus heparin			2	5
Knee length stocking plus heparin				1

was the most effective treatment with an absolute risk of DVT of 4.04%. The probability that thigh length stockings with pharmacological prophylaxis is the most effective treatment in a new trial of all the treatments is 73%, as displayed in table 2. The probability of being the most effective treatment does not simply reflect the effectiveness of the treatment, but also the uncertainty in the estimate. While thigh length stockings plus pharmacological prophylaxis appear to be the most effective treatment with little uncertainty, the marginal benefit of thigh length stockings plus heparin over heparin alone is less than the marginal benefit of heparin over no treatment as heparin has already reduced the risk of DVT substantially.

The sensitivity analysis modelling an interaction between thigh or knee length stockings and heparin produced results with the same direction of effect but greater uncertainty in the effect estimate (knee vs thigh OR 2.59, CrI 0.92 to 7.84,  $p=0.10$ ). The subgroup analysis suggested that thigh length stockings with heparin appear to be more effective in the non-orthopaedic

**Table 2** Probability of being the most effective treatment in a new trial of all treatments

Treatment	Probability of being the most effective treatment	Absolute risk of DVT (%)
No treatment	0.00	29.28
Thigh length stocking	0.04	13.76
Knee length stocking	0.02	22.01
Heparin	0.02	9.88
Thigh length stocking plus heparin	0.73	4.04
Knee length stocking plus heparin	0.20	6.94

DVT, deep vein thrombosis.

surgery group than in the orthopaedic surgery group. The median ORs are slightly more in favour of both thigh and knee length stockings with heparin compared to heparin alone for the non-orthopaedic surgery group (thigh: median OR 3.83, 95% CrI 2.29 to 6.66,  $p=0.00$ ; knee: median OR 2.16, 95% CrI 0.90 to 5.21,  $p=0.09$ ) compared to the orthopaedic surgery group (thigh: median OR 2.05, 95% CrI 1.32 to 3.23,  $p=0.00$ ; knee: median OR 1.32, 95% CrI 0.72 to 2.46,  $p=0.37$ ).

### PE, mortality and adverse event results

Fifteen RCTs assessed PE or fatal PE, 11 RCTs assessed mortality and 12 RCTs reported results relating to adverse events. PE events and VTE-related deaths were generally rare in the included trials. Adverse events were rarely reported and those related to antiembolism stockings were minor events, including minor foot abrasions, superficial thrombophlebitis or the stocking slipping down. The majority of complications reported were minor bleeding complications associated with pharmacoprophylaxis, although the proportion of patients reporting such events was low; between 1% and 4%.

### DISCUSSION

This systematic review assessed the clinical effectiveness of thigh length versus knee length antiembolism stockings for the prevention of DVT in surgical patients. The review only included studies of surgical patients; therefore, the results are not generalisable to other patient populations, who may have a different baseline risk of DVT and of the adverse effects of thromboprophylaxis. Patients with stroke have been investigated separately in a large RCT of thigh length versus knee length antiembolism stockings, which found that DVT occurred more often in patients who wore knee length stockings than those who wore thigh length stockings.<sup>36</sup>

A previous Cochrane review comparing knee length versus thigh length antiembolism stockings in post-operative surgical patients included three of the five RCTs included in our direct meta-analysis.<sup>11</sup> The Cochrane review also found no statistically significant difference in clinical effectiveness between the two stocking lengths in terms of reducing the incidence of DVT. The authors concluded that there was insufficient high quality evidence to determine whether thigh length or knee length stockings differ in their effectiveness in terms of reducing the incidence of DVT in hospital in patients.

Our systematic review included a network meta-analysis of all the trials that indirectly informed the relative effectiveness of thigh length versus knee length antiembolism stockings, with or without pharmacological prophylaxis, for the prevention of DVT in surgical patients. The results of the NMA were consistent with the direct meta-analysis, without increasing the precision of the estimates. Overall, thigh length stockings with pharmacological prophylaxis appears to be the most

effective method of preventing DVT in surgical patients, the NMA results also indicate that the marginal benefit of thigh length stockings plus heparin over heparin alone is less than the marginal benefit of heparin over no treatment, as heparin already reduces the risk of DVT substantially.

Evidence relating to other outcomes was sparse; few trials reported complications and consequences associated with DVT, such as the incidence of PE, post-thrombotic syndrome and mortality or adverse effects.

Despite the weak evidence base and importance of the question, it is unlikely to be worthwhile undertaking a new definitive trial comparing thigh length versus knee length antiembolism stockings. Such a trial would need to be very large to enable assessment of clinically relevant DVT and its associated complications and consequences in the relevant population, and should include an assessment of patient adherence, both in hospital and after patients have been discharged home. Such a trial would therefore, be very costly to run. In addition, while thigh length stockings appear to have superior efficacy, practical issues may prevent their wide use in clinical practice; patients report that both thigh length and knee length stockings are difficult to use, but fewer patients report discomfort with knee length stockings and patients are more likely to wear knee length stockings correctly.<sup>3-5</sup> A more pragmatic approach may be to give thigh length stockings only to patients who can use them properly and consistently, while knee length stockings are more appropriate for others.

### Limitations

There was substantial variation across the included trials in terms of the patient characteristics (suggesting that the participants had a different baseline risk for DVT) and interventions used (in terms of both stocking use and concomitant pharmacological prophylaxis). The timing of outcome assessments was generally short, where reported; therefore some DVTs may have been missed. The included trials assessed all DVTs, not just symptomatic DVTs; where reported the majority of DVTs were asymptomatic, the clinical consequences of which are unknown.

Many of the included trials dated back to the 1970s and 1980s, therefore, they may not reflect current practice: surgical practice has changed over time with less invasive surgical procedures, shorter duration of hospitalisation and earlier mobilisation after surgery.

Generally the trial methods were poorly reported, making risk of bias assessment difficult. Only three out of 23 included RCTs were considered to have a low risk of bias; the reporting was inadequate to judge the risk of bias for most trials. This systematic review included all relevant trials, regardless of trial quality; therefore, the uncertain quality of many of the included trials reduces the reliability of the results of this review.

### Conclusions

The evidence base for assessing the relative treatment effectiveness of thigh length and knee length antiembolism stockings for the prevention of DVT in surgical patients is weak; most studies are old and may not reflect current practice.

However, direct and indirect meta-analysis suggests that thigh length stockings may be more effective than knee length stockings, although the results were not statistically significant. Overall, thigh length stockings with pharmacological prophylaxis appears to be the most effective method of preventing DVT in surgical patients, although the marginal benefit of thigh-length stockings plus heparin over heparin alone is less than the marginal benefit of heparin over no treatment as heparin already reduces the risk of DVT substantially.

### Recommendations

Thigh length antiembolism stockings may be more effective than knee length stockings at DVT prevention in surgical patients; however, much of the available research evidence is old and of uncertain quality. A definitive trial in high risk surgical patients to compare thigh length versus knee length antiembolism stockings, in addition to standard pharmacological prophylaxis, would need to be very large to enable assessment of clinically relevant DVT and its associated complications and consequences. Therefore, such a trial would be very costly to run and it is not clear that it would be worthwhile. A more pragmatic approach may be to give thigh length stockings only to patients who can use them properly and consistently, while knee length stockings are more appropriate for patients who are less physically adept or likely to be less compliant.

### Author affiliations

<sup>1</sup>Centre for Reviews and Dissemination, University of York, York, UK

<sup>2</sup>Northern Vascular Centre, Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK

<sup>3</sup>Department of Spinal Surgery, Leeds Teaching Hospitals NHS Trust, Leeds, UK

<sup>4</sup>Department of Pathology, Royal Bournemouth and Christchurch Hospitals NHS Foundation Trust, Bournemouth, UK

**Acknowledgements** The authors would like to thank our patient representative, Mrs Kaye Norman, for advice.

**Contributors** RW wrote the protocol, performed study selection, data extraction, quality assessment and synthesis of the included studies and drafted this paper. FP contributed to protocol development, performed study selection, data extraction, quality assessment and synthesis of the included studies and commented on this paper. SR undertook the network meta-analysis and commented on this paper. GS provided clinical advice throughout the project and commented on this paper. PM and HF provided clinical advice throughout the project and commented on this paper. DF developed and conducted the search strategies and commented on this paper. NW contributed to protocol development, undertook an advisory role and took overall managerial responsibility for the project, as well as commenting on this paper.

**Funding** This paper summarises a systematic review undertaken as part of a larger research project commissioned by the National Institute for Health Research (NIHR) Health Technology Assessment Programme as project

number HTA 13/72/01. The views expressed are those of the authors and not necessarily those of the NIHR HTA Programme.

**Competing interests** None declared.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data sharing statement** Additional data are presented in a Health Technology Assessment report.

**Open Access** This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY 4.0) license, which permits others to distribute, remix, adapt and build upon this work, for commercial use, provided the original work is properly cited. See: <http://creativecommons.org/licenses/by/4.0/>

## REFERENCES

- House of Commons Health Committee. *The prevention of venous thromboembolism in hospitalised patients: second report of session 2004–05*. London: The Stationery Office Limited, 2005.
- Scottish Intercollegiate Guidelines Network (SIGN). *SIGN guideline 122: prevention and management of venous thromboembolism: a national clinical guideline*. Edinburgh: Scottish Intercollegiate Guidelines Network, 2010.
- Brady D, Raingruber B, Peterson J, *et al*. The use of knee-length versus thigh-length graduated compression stockings and sequential compression devices. *Crit Care Nurs Q* 2007;30:255–62.
- Hameed MF, Browse DJ, Immelman EJ, *et al*. Should knee-length replace thigh-length graduated compression stockings in the prevention of deep-vein thrombosis? *S Afr J Surg* 2002;40:15–16.
- Benkő T, Cooke EA, McNally MA, *et al*. Graduated compression stockings: knee length or thigh length. *Clin Orthop Relat Res* 2001;383:197–203.
- National Clinical Guideline Centre - Acute and Chronic Conditions. *NICE clinical guidance CG92: venous thromboembolism: reducing the risk: reducing the risk of venous thromboembolism (deep vein thrombosis and pulmonary embolism) in patients admitted to hospital*. London: National Institute for Health and Clinical Excellence, 2010.
- Wade R, Sideris E, Paton F, *et al*. Graduated compression stockings for the prevention of deep-vein thrombosis in postoperative surgical patients: a systematic review and economic model with a value of information analysis. *Health Technol Assess* 2015;19:1–220.
- Higgins JPT, Green S. *Cochrane handbook for systematic reviews of interventions version 5.1.0 [updated March 2011]*. Oxford: The Cochrane Collaboration, 2011.
- Dias S, Welton N, Sutton A, *et al*. *NICE DSU technical support document 2: a generalised linear modelling framework for pairwise and network meta-analysis of randomised controlled trials*. Sheffield: Decision Support Unit, SchARR, University of Sheffield, 2011 (updated 2014).
- Sachdeva A, Dalton M, Amaragiri SV, *et al*. Elastic compression stockings for prevention of deep vein thrombosis. *Cochrane Database Syst Rev* 2010;(7):CD001484.
- Sajid MS, Desai M, Morris RW, *et al*. Knee length versus thigh length graduated compression stockings for prevention of deep vein thrombosis in postoperative surgical patients. *Cochrane Database Syst Rev* 2012;5:CD007162.
- Howard A, Zaccagnini D, Ellis M, *et al*. Randomized clinical trial of low molecular weight heparin with thigh-length or knee-length antiembolism stockings for patients undergoing surgery. *Br J Surg* 2004;91:842–7.
- Barnes RW, Brand RA, Clarke W, *et al*. Efficacy of graded-compression anti-embolism stockings in patients undergoing total hip arthroplasty. *Clin Orthop Relat Res* 1978;132:61–7.
- Bergqvist D, Lindblad B. The thromboprophylactic effect of graded elastic compression stockings in combination with Dextran 70. *Arch Surg* 1984;119:1329–31.
- Kalodiki EP, Hoppensteadt DA, Nicolaidis AN, *et al*. Deep venous thrombosis prophylaxis with low molecular weight heparin and elastic compression in patients having total hip replacement—a randomised controlled trial. *Int Angiol* 1996;15:162–8.
- Scurr JH, Coleridge-Smith PD, Hasty JH. Regimen for improved effectiveness of intermittent pneumatic compression in deep venous thrombosis prophylaxis. *Surgery* 1987;102:816–20.
- Shirai N. Study on prophylaxis of postoperative deep vein thrombosis. *Acta Scholae Medicinalis Universitatis in Gifu* 1985;33:1173–83.
- Tömgren S. Low dose heparin and compression stockings in the prevention of postoperative deep venous thrombosis. *Br J Surg* 1980;67:482–4.
- Tsapogas MJ, Goussous H, Peabody RA, *et al*. Postoperative venous thrombosis and effectiveness of prophylactic measures. *Arch Surg* 1971;103:561.
- Turpie AGG, Hirsh J, Gent M, *et al*. Prevention of deep-vein thrombosis in potential neurosurgical patients—a randomized trial comparing graduated compression stockings alone or graduated compression stockings plus intermittent pneumatic compression with control. *Arch Intern Med* 1989;149:679–81.
- Wille-jørgensen P, Hauch O, Dimo B, *et al*. Prophylaxis of deep venous thrombosis after acute abdominal operation. *Surg Gynecol Obstet* 1991;172:44–8.
- Wille-jørgensen P, Thorup J, Fischer A, *et al*. Heparin with and without graded compression stockings in the prevention of thromboembolic complications of major abdominal surgery: a randomized trial. *Br J Surg* 1985;72:579–81.
- Camporese G, Bernardi E, Prandoni P, *et al*. Low-molecular-weight heparin versus compression stockings for thromboprophylaxis after knee arthroscopy: a randomized trial. *Ann Intern Med* 2008;149:73–82.
- Caprini JA, Chucker JL, Zuckerman L, *et al*. Thrombosis prophylaxis using external compression. *Surg Gynecol Obstet* 1983;156:599–604.
- Cohen AT, Skinner JA, Warwick D, *et al*. The use of graduated compression stockings in association with fondaparinux in surgery of the hip: a multicentre, multinational, randomised, open-label, parallel-group comparative study. *J Bone Joint Surg Br* 2007;89B:887–92.
- Fasting H, Andersen K, Nielsen HK, *et al*. Prevention of postoperative deep venous thrombosis—low-dose heparin versus graded pressure stockings. *Acta Chir Scand* 1985;151:245–8.
- Ishak MA, Morley KD. Deep venous thrombosis after total hip-arthroplasty: a prospective controlled-study to determine the prophylactic effect of graded pressure stockings. *Br J Surg* 1981;68:429–32.
- Mellbring G, Palmer K. Prophylaxis of deep-vein thrombosis after major abdominal-surgery—comparison between dihydroergotamine-heparin and intermittent pneumatic calf compression and evaluation of added graduated static compression. *Acta Chir Scand* 1986;152:597–600.
- Porteous MJL, Nicholson EA, Morris LT, *et al*. Thigh length versus knee length stockings in the prevention of deep vein thrombosis. *Br J Surg* 1989;76:296–7.
- Rasmussen A, Hansen PT, Lindholt J, *et al*. Venous thrombosis after abdominal-surgery—a comparison between subcutaneous heparin and antithrombotic stockings, or both. *J Med* 1988;19:193–201.
- Fredin H, Bergqvist D, Cederholm C, *et al*. Thromboprophylaxis in hip-arthroplasty: dextran with graded compression or preoperative dextran compared in 150 patients. *Acta Orthop Scand* 1989;60:678–81.
- Hui ACW, HerasPalou C, Dunn I, *et al*. Graded compression stockings for prevention of deep-vein thrombosis after hip and knee replacement. *J Bone Joint Surg Br* 1996;78B:550–4.
- Williams JT, Palfrey SM. Cost effectiveness and efficacy of below knee against above knee graduated compression stockings in the prevention of deep vein thrombosis. *Phlebologie* 1988;41:809–11.
- Ayhan H, Iyigun E, Ince S, *et al*. Comparison of three different protocols in the prevention of postoperative deep vein thrombosis in patients at high-risk: randomized clinical study. *Eur Surg Res* 2013;50:64–5.
- Kahn SR, Lim W, Dunn AS, *et al*. Prevention of VTE in nonsurgical patients: antithrombotic therapy and prevention of thrombosis, 9th ed: American College of Chest Physicians evidence-based clinical practice guidelines. *Chest* 2012;141: e195S–226S.
- The CLOTS Trial Collaboration. High-length versus below-knee stockings for deep venous thrombosis prophylaxis after stroke: a randomized trial. *Ann Intern Med* 2010; 153:553–62.

## REVIEW PAPER

# Systematic review of patient preference and adherence to the correct use of graduated compression stockings to prevent deep vein thrombosis in surgical patients

Ros Wade, Fiona Paton & Nerys Woolacott

Accepted for publication 26 July 2016

Correspondence to R. Wade:  
e-mail: ros.wade@york.ac.uk

Ros Wade MSc  
Research Fellow  
Centre for Reviews and Dissemination,  
University of York, UK

Fiona Paton MSc  
Research Fellow  
Centre for Reviews and Dissemination,  
University of York, UK

Nerys Woolacott BSc PhD  
Senior Research Fellow  
Centre for Reviews and Dissemination,  
University of York, UK

WADE R., PATON F. & WOOLACOTT N. (2017) Systematic review of patient preference and adherence to the correct use of graduated compression stockings to prevent deep vein thrombosis in surgical patients. *Journal of Advanced Nursing* 73(2), 336–348. doi: 10.1111/jan.13148

## Abstract

**Aim.** The aim of this study was to explore patient preference and adherence to thigh and knee length graduated compression stockings for the prevention of deep vein thrombosis in surgical patients.

**Background.** Hospitalised patients are at risk of developing deep vein thrombosis. Mechanical methods of prophylaxis include compression stockings, available as knee or thigh length. Patient adherence to correct stocking use is of critical importance to their effectiveness.

**Design.** Systematic review of quantitative evidence.

**Data sources.** Eleven databases were searched from inception to 2013 for systematic reviews of compression stockings. Reviews were screened for relevant primary studies and update searches of eight electronic sources were undertaken (2010–2014).

**Review methods.** Randomised controlled trials and observational studies of surgical patients using compression stockings were quality assessed and data were extracted on patient adherence and preference. A narrative summary is presented.

**Results.** Nine randomised controlled trials and seven observational studies were included in the systematic review. There was substantial variation between studies in terms of patient characteristics, interventions and methods of outcome assessment.

**Conclusion.** Patient adherence was generally higher with knee length than thigh length stockings. However, the studies reflect patient adherence in a hospital setting only, where patients are observed by healthcare professionals; it is likely that adherence reduces once patients have been discharged from hospital. Patients preferred knee length stockings over thigh length stockings. In many clinical settings, any difference in efficacy between thigh length and knee length stockings may be rendered irrelevant by patient preference for and likely better adherence to knee length stockings.

**Keywords:** anti-embolism stocking, deep vein thrombosis, graduated compression stocking, literature review, nursing, patient adherence, patient preference, systematic review

### Why is this review needed?

- Surgical patients are at an increased risk of developing deep vein thrombosis. National Institute for Health and Care Excellence clinical guidelines recommend various forms of prophylaxis, including mechanical methods, such as graduated compression stockings.
- Compression stockings are available as knee or thigh length; patients report difficulties with the use of both stocking lengths. The incorrect use of compression stockings can be unsafe.
- Patient adherence is of critical importance to the effectiveness of compression stockings for the prevention of deep vein thrombosis.

### What are the key findings?

- There is substantial variation in the characteristics of studies assessing patient preference and adherence to compression stockings, many studies are poorly reported with an unclear risk of bias.
- Patient adherence was generally higher with knee length than thigh length stockings.
- Patients preferred knee length stockings over thigh length stockings.

### How should the findings be used to influence policy/practice/research?

- Efforts need to be made to improve patient adherence to the correct use of compression stockings, particularly of thigh length stockings.
- The choice between thigh and knee length stockings should take into account the likely adherence given each individual patient's particular needs and circumstances.
- Any future research into the effectiveness of compression stockings should take into account patient preference and incorporate assessment of patient adherence, both in a hospital setting and post discharge.

## Introduction

Venous thrombosis is a condition where a blood clot forms in a vein, most commonly the deep veins of the legs. The clot may break off and travel to the lungs, causing a potentially fatal pulmonary embolism (PE). Deep vein thrombosis (DVT) is usually asymptomatic and only detected during screening, although it may be associated with leg pain and/or swelling as a result of occlusion of the vein (National Institute for Health and Care Excellence 2010).

Certain people, particularly hospitalised patients, are at greater risk of DVT. Factors such as immobilisation,

decreased fluid intake and excessive body fluid loss may cause changes in the blood vessel wall, blood flow and properties of the blood. Trauma and surgery can also increase the risk of DVT (Sajid *et al.* 2012).

In the UK, National Institute for Health and Care Excellence (NICE) clinical guidelines recommend various forms of prophylaxis, based on their effectiveness to reduce the risk of DVT and taking into account individual patient factors and according to clinical judgement. Prophylaxis can be pharmacological (usually low molecular weight heparin or unfractionated heparin) and/or mechanical. Mechanical methods of prophylaxis include graduated compression stockings, intermittent pneumatic compression devices and pneumatic foot pumps. Graduated compression stockings exert pressure at a decreasing gradient from the ankle towards the thigh, which increases blood flow and promotes venous return, thus preventing passive venous distension which is thought to prevent sub-endothelial tears and activation of clotting factors (National Institute for Health and Care Excellence 2010). Evidence suggests that compression stockings can reduce the incidence of postoperative DVT to approximately 11%, while low-dose heparin reduces the rate to approximately 9%; used together the rate of DVT is reduced further (Nicolaidis *et al.* 2001).

## Background

Graduated compression stockings are available as knee or thigh length stockings. A recent systematic review identified that thigh length stockings are probably more effective than knee length stockings at preventing DVT in surgical patients, though because of limitations there is some uncertainty about this, in particular how they perform in the real world setting (Wade *et al.* 2015). Patients report that both stocking lengths are difficult to use (Hameed *et al.* 2002, Brady *et al.* 2007). The incorrect use of compression stockings can be unsafe: thigh length stockings that are fitted incorrectly or that roll down the leg can create a tourniquet effect, which can potentially damage skin and reduce venous outflow (Sajid *et al.* 2012). In addition, one length of stocking may be more appropriate than the other in certain patients; knee length stockings may be more likely to induce wound complications in patients undergoing knee replacement surgery as the elastic support lies over the wound, creating unwanted localised pressure. There are also some patients for whom compression stockings are contraindicated, such as those who have peripheral arterial disease, cardiac failure and severe leg oedema (National Institute for Health and Care Excellence 2010).

In practice, patients' ability and willingness to wear stockings correctly is of critical importance to their effectiveness. Non-adherence to interventions, defined as the extent to which the patient's action does not match the prescriber's agreed recommendations (National Institute for Health and Care Excellence 2009), may reduce the benefits of interventions on health. Non-adherence may be the result of patient behaviour, but it could also reflect a fundamental limitation in the delivery of health care. In this specific case, for example, there may be a failure to correctly educate patients in the use of knee or thigh length stockings, lack of monitoring of correct usage or a failure to identify and provide the support that patients need post discharge (National Institute for Health and Care Excellence 2010). An earlier systematic review conducted for NICE guidelines identified only one randomised controlled trial (RCT) and two observational studies assessing patient views and adherence to mechanical devices (National Institute for Health and Care Excellence 2010).

## The review

### Aim

The aim of this systematic review was to assess the evidence on patient adherence and preference for knee or thigh length graduated compression stockings for the prevention of DVT in surgical patients. This review was undertaken as part of a larger project comparing the clinical and cost-effectiveness of knee and thigh length stockings.

### Design

The systematic review was conducted and reported following the general principles for conducting a systematic review of health interventions recommended in CRD's guidance for undertaking reviews in health care (Centre for Reviews and Dissemination 2009) and the reporting guidance of the PRISMA statement (Moher *et al.* 2009). The research protocol for the broader project of which this systematic review is a part, was registered on the international prospective register of systematic reviews (PROSPERO registration number: CRD42014007202).

### Search methods

A systematic approach to identifying the evidence on patient preference and/or adherence with regard to graduated compression stockings was undertaken. The literature search was conducted as part of the broader project and

encompassed all publications relevant to graduated compression stockings in surgical patients (Wade *et al.* 2015). In the first instance, existing systematic reviews were sought to identify relevant primary studies. Eleven guideline and systematic review databases (including the Cochrane Database of Systematic Reviews, Database of Abstracts of Reviews of Effects, PROSPERO, Health Technology Assessment Database and National Guidelines Clearinghouse) were searched from inception to August 2013. The search was then brought up to date in February 2014 using systematic searches of six electronic databases (MEDLINE, MEDLINE In-Process, EMBASE, CINAHL, AMED and CENTRAL) and two grey literature databases (ClinicalTrials.gov and Current Controlled Trials) to identify RCTs published after January 2010 (January 2010 - February 2014). No language restrictions were applied to the search strategies. The search strategy developed for Ovid MEDLINE is presented in supplementary Table S1.

Eligible for inclusion in the review were trials and observational studies of patients undergoing surgery (both day surgery and inpatients). Eligible studies had to assess thigh length vs. knee length graduated compression stockings (with or without standard pharmacological prevention) or compression stockings vs. no stocking. The outcome of interest was patient adherence to wearing knee length or thigh length stockings and patient preference in terms of length of stocking. The inclusion criteria are stated below in PICO format.

*Participants:* Patients undergoing surgery (both day surgery and inpatients).

*Intervention:* Thigh length or knee length graduated compression stockings (with or without standard pharmacological prophylaxis).

*Comparison:* The alternative length of stocking or no stocking.

*Outcomes:* Patient adherence to wearing knee length or thigh length stockings and patient preference in terms of length of stocking.

### Search outcomes

Search results were exported into Endnote® Version 7.2. Two reviewers independently screened records for inclusion. The findings from studies identified from the search undertaken as part of the broader review were consistent in terms of patient experiences in wearing the two different lengths of stocking. A separate search of the literature was therefore not undertaken to identify further relevant observational studies as it was deemed that additional evidence would not have substantially added to the evidence base.

## Quality appraisal

RCTs were quality assessed according to the Cochrane Risk of Bias Tool (Higgins & Green 2011). The observational studies were assessed on whether they fulfilled the following criteria: prospective design; matched control group; consecutive recruitment of patients; clear description of stockings; clear description of patients. A more formal assessment of study quality was not deemed appropriate given the nature of these studies.

## Data abstraction

Data on patient adherence and preference reported in RCTs were extracted into EPPI-Reviewer 4.0 by one reviewer. Data from the observational studies were extracted into a Microsoft Word document by one reviewer. A second reviewer checked all data for accuracy.

## Synthesis

Given the heterogeneity between the studies and the limited quality of the studies and limited amount of outcome data reported, data are presented in tables and as a narrative summary.

## Results

The review identified nine RCTs (Wille-Jorgensen *et al.* 1985, Mellbring & Palmer 1986, Fredin *et al.* 1989, Porteous *et al.* 1989, Turpie *et al.* 1989, Hui *et al.* 1996, Benko *et al.* 2001, Camporese *et al.* 2008, Ayhan *et al.* 2013) and seven observational studies (reported in eight articles) (Williams *et al.* 1994, 1996, Hameed *et al.* 2002, Parnaby 2004, Williams & Owen 2006, Brady *et al.* 2007, Winslow & Brosz 2008, Thompson *et al.* 2011) that provided data on patient adherence and/or preference (Tables 1 & 2 for study characteristics). Figure 1 presents the flow of studies through the study selection process.

Most of the included studies were conducted in the UK or Europe and all of the studies were reported in English. Studies were published between 1985 - 2013; many of the RCTs dated back to the 1980s, therefore, their results may not be generalisable to current practice. Four of the RCTs included patients undergoing orthopaedic surgery (Fredin *et al.* 1989, Hui *et al.* 1996, Benko *et al.* 2001, Camporese *et al.* 2008), three included patients undergoing abdominal surgery (Wille-Jorgensen *et al.* 1985, Mellbring & Palmer 1986, Porteous *et al.* 1989), one included patients undergoing neurosurgery (Turpie *et al.* 1989) and one RCT did not state the type of surgery (Ayhan *et al.* 2013). Three of the

observational studies included patients undergoing orthopaedic surgery (Williams *et al.* 1996, Williams & Owen 2006, Thompson *et al.* 2011), three included patients from mixed surgical units (Hameed *et al.* 2002, Parnaby 2004, Winslow & Brosz 2008) and one observational study included patients admitted to a range of acute care nursing units (Brady *et al.* 2007). Sample sizes ranged from 114-1761 in the RCTs and from 50-324 in the observational studies.

Methods for measuring adherence were unclear in some studies and definitions for this outcome were inconsistent across RCTs and observational studies (see Table 3). For example, some studies measured stocking removal/treatment discontinuation, others measured tolerance to stockings or correct usage/management of stockings. Studies also varied in terms of characteristics and methodology, which may account for some of the differences in levels of patient adherence and preference for one length of stocking over the other.

Results from the quality assessment of RCTs are presented in supplementary Table S2. One RCT was considered to be at low risk of bias (Camporese *et al.* 2008), two RCTs were considered to be at high risk of bias (Wille-Jorgensen *et al.* 1985, Hui *et al.* 1996) and the risk of bias could not be determined for the remaining six RCTs due to poor reporting of study methods (Mellbring & Palmer 1986, Fredin *et al.* 1989, Porteous *et al.* 1989, Turpie *et al.* 1989, Benko *et al.* 2001, Ayhan *et al.* 2013).

The observational studies were generally small surveys of the usage of graduated compression stockings (mostly <150 patients) and most studies did not use rigorous methods (see supplementary Table S3). All were prospective studies, but none included matched control groups. In most cases, the patients were a convenience sample, of those who had been admitted for surgery or were on the selected hospital ward. Other than the broad label of the surgical procedure undertaken, most studies failed to provide demographic details of the patients studied. The specific brands of the stockings studied were often not reported, although there was a clear distinction made between thigh length and knee length stockings in the reporting of the studies.

## Patient adherence

Seven RCTs and six observational studies reported data on patient adherence (Table 3) (Wille-Jorgensen *et al.* 1985, Mellbring & Palmer 1986, Fredin *et al.* 1989, Turpie *et al.* 1989, Hui *et al.* 1996, Williams *et al.* 1996, Benko *et al.* 2001, Hameed *et al.* 2002, Parnaby 2004, Brady *et al.* 2007, Camporese *et al.* 2008, Winslow & Brosz 2008, Thompson *et al.* 2011). Two RCTs compared thigh length

Table 1 Study characteristics for RCTs.

Author	Location and number randomised/analysed	Patient characteristics	Type of surgery	Intervention	Control
Ayhan <i>et al.</i> (2013)	Turkey 219/NR	VTE risk factors: Extremely high risk for postoperative DVT (not defined) Average age: NR Male: NR	NR	Low pressure thigh length GCS ( <i>n</i> = not stated)	(1) Low pressure knee length GCS ( <i>n</i> = not stated) (2) Moderate pressure knee length GCS ( <i>n</i> = not stated) No stocking ( <i>n</i> = 40)
Benko <i>et al.</i> (2001)	Belfast 200/200	VTE risk factors: 94/200 (47%) had varicose veins Average age: 60.4 years Male: 45%	Orthopaedic	(1) Thrombex anti-embolism stocking ( <i>n</i> = 40) (2) Brevet TX anti-embolism stocking ( <i>n</i> = 40)	
Camporese <i>et al.</i> (2008)	Italy 1,761/1,602	VTE risk factors: 26% were smokers, 1% had family history of VTE, 9% were using hormonal compounds and the average body mass index was around 25.5. Patients with previous VTE or active malignancy were excluded from the trial Average age: 42 years Male: ratio around 1.6:1	Orthopaedic	Thigh length GCS (pressure at the ankle 30-40 mm Hg) worn on the operated on leg ( <i>n</i> = 597)	Nadroparin LMWH for 7 days ( <i>n</i> = 603)
Fredin <i>et al.</i> (1989)	Sweden 150/144	VTE risk factors: NR Average age: 67 years Male: 40%	Orthopaedic	Comprinert thigh length GCS plus dextran ( <i>n</i> = 49)	Nadroparin LMWH for 14 days ( <i>n</i> = 402). This arm was stopped early for safety reasons (1) Dextran ( <i>n</i> = 48) (2) Dextran plus additional preoperative dextran ( <i>n</i> = 47) (1) Knee length GCS ( <i>n</i> = 40)
Hui <i>et al.</i> (1996)	UK 177/138	VTE risk factors: 14% were smokers, average body mass index was nearly 28 for patients undergoing total knee replacement Average age: 69 years (range 49-88) Male: NR	Orthopaedic	Thigh length GCS ( <i>n</i> = 44)	(2) No stocking ( <i>n</i> = 54)
Mellbring and Palmer (1986)	Sweden 114/108	VTE risk factors: 36% had malignant disease and the average body mass index was 25.3 Average age: 66 years (range 50-85) Male: 47%	Abdominal	Low dose heparin plus dihydro-ergotamine ( <i>n</i> = 54)	Intermittent pneumatic calf compression ( <i>n</i> = 54)
Porteous <i>et al.</i> (1989)	UK 124/114	VTE risk factors: 40% had malignancy, 32% were smokers. Patients with a history of DVT were excluded from the trial Average age: 65 years Male: 43%	Abdominal	Each patient wore a thigh length GCS randomly allocated to the right or left leg Thigh length GCS (pressure at the ankle: 11.3 mm Hg) ( <i>n</i> = 56)	Each patient wore a thigh length GCS randomly allocated to the right or left leg Knee length GCS ( <i>n</i> = 58)

**Table 1** (Continued).

Author	Location and number randomised/analysed	Patient characteristics	Type of surgery	Intervention	Control
Turpie <i>et al.</i> (1989)	Canada 239/239	VTE risk factors: 3% had a history of DVT, 49% had a brain or spinal cord tumour and 23% had subarachnoid haemorrhage Average age: 50 years (range 16-90) Male: 60%	Neurosurgery	(1) Kendall thigh length GCS ( <i>n</i> = 80)  (2) Kendall thigh length GCS plus intermittent pneumatic compression ( <i>n</i> = 78) Thigh length GCS plus low dose heparin ( <i>n</i> = 86)	No treatment ( <i>n</i> = 81)
Wille-Jorgensen <i>et al.</i> (1985)	Denmark 196/176	VTE risk factors: All patients had at least one thromboembolic risk factor Average age: 60 years (range 36-90) Male: 60%	Abdominal		Low dose heparin ( <i>n</i> = 90)

DVT, deep vein thrombosis; GCS, graduated compression stockings; LMWH, low molecular weight heparin; NR, not reported; VTE, venous thromboembolism.

vs. knee length stockings. Hui *et al.* (1996) included 138 patients undergoing total hip or knee replacement and reported that a higher proportion of patients wearing thigh length compared with knee length compression stockings removed the stockings due to discomfort (23% vs. 16% respectively). Benko *et al.* (2001) assessed 200 patients undergoing orthopaedic surgery, after one hour, statistically significantly more wrinkles and discomfort were reported by patients wearing thigh length stockings. Fifty per cent of patients were unable to fit the stockings independently, with similar numbers of patients distributed between thigh length and knee length stocking groups (Benko *et al.* 2001).

Three RCTs comparing thigh length stockings plus other treatment vs. other treatment alone provided limited details on patient adherence. Fredin *et al.* (1989) compared thigh length stockings plus dextran vs. dextran alone in 144 orthopaedic patients, reporting that two of 49 (4%) patients in the stockings plus dextran group discontinued wearing stockings because of discomfort. Wille-Jorgensen *et al.* (1985) reported that two of 86 patients (2.3%) undergoing abdominal surgery removed their thigh length stockings after 5 days, but they were otherwise well tolerated. The authors did not report the reasons for removal of the stockings (Wille-Jorgensen *et al.* 1985). Mellbring and Palmer (1986) simply stated that all patients undergoing abdominal surgery (*n* = 108) tolerated wearing thigh length stockings.

The other two RCTs compared compression stockings with another method of thromboprophylaxis. Camporese *et al.* (2008) compared thigh length stockings with two different regimens of low molecular weight heparin (LMWH); similar proportions of patients from each group declined to complete the prophylactic regimen. Turpie *et al.* (1989) reported that 2.5% patients did not wear thigh length stockings correctly.

All of the observational studies were conducted in a hospital setting and most patients were assigned to wear thigh length stockings. Across the six studies that reported on adherence, this was relatively poor (Williams *et al.* 1996, Hameed *et al.* 2002, Parnaby 2004, Brady *et al.* 2007, Winslow & Brosz 2008, Thompson *et al.* 2011). The proportion of patients not wearing their stockings or wearing them incorrectly appeared to be generally higher in patients receiving thigh length stockings (Table 3). Although the objectives of the observational studies were specifically to assess the correct use of thigh length and knee length compression stockings and to elicit patient perspectives about their use, they reflect adherence only in a hospital setting where patients are observed by healthcare professionals.

Across the studies, reasons for not wearing stockings were related to discomfort, stocking provision, removing stockings for bathing or no longer requiring them due to

**Table 2** Study characteristics for observational studies.

Author	Study design	Location and number recruited	Patient characteristics	Type of surgery	Intervention
Brady <i>et al.</i> (2007)	Single centre survey	USA 137	Average age: NR (range 18-92 years) Male: 47%	Mixed	Thigh length or knee length GCS and/or sequential compression device
Hameed <i>et al.</i> (2002)	Single centre prospective observational study	South Africa 72	Average age: 51 years (range 13-84) Male: 54%	Mixed	Kendall thigh length or knee length GCS
Parnaby (2004)	Single centre survey and two trials	UK 218 (survey); 70 (trial 1); 20 (trial 2)	Average age: NR Male: NR	Mixed	Thigh length or knee length SaphenaMedical anti-DVT GCS
Thompson <i>et al.</i> (2011)	Audit and trial	UK 62 (57 analysed)	Average age: NR Male: 44%	Orthopaedic	Knee length Preventex GCS
Williams <i>et al.</i> (1994, 1996)	Multicentre prospective observational study	UK N = 324 (131 wore stockings)	Average age: 67 years Male: 40%	Orthopaedic	Thigh length (brand unknown) or Brevet GCS, or knee length Brevet GCS
Williams and Owen (2006)	Single centre prospective observational study	UK 50 (47 analysed)	Average age: 71 years Male: 53%	Orthopaedic	Thigh length or knee length GCS
Winslow and Brosz (2008)	Single centre prospective observational study	USA 145 (142 analysed)	Average age: 57 years (range 18-97) Male: 16%	Mixed	Thigh length or knee length GCS

DVT, deep vein thrombosis; GCS, graduated compression stockings; NR, not reported.

ambulation. Incorrect use related to wearing incorrectly sized stockings, or rolling down, binding, or wrinkling of the stocking. In a study involving an audit and a trial of knee length compression stockings, 74% of patients were wearing stockings that were incorrectly sized (Thompson *et al.* 2011). Implementation of a standardised protocol in the study setting resulted in statistically significant reductions in this proportion to 34%; this figure still appears high (Thompson *et al.* 2011).

Brady *et al.* (2007) reported a strong correlation between age and adherence, indicating that older patients wore stockings more consistently compared to younger patients (Pearson correlation = 0.247;  $P = 0.01$  [no confidence intervals presented]). None of the other studies assessed this association and the evidence is therefore insufficient to draw any conclusions.

Three observational studies directly compared adherence to thigh length vs. knee length stockings. Non-adherence (not wearing the stocking at all) was worse with thigh length stockings than knee length stockings (16.7% vs. 3%, respectively) (Hameed *et al.* 2002). Incorrect usage was also higher with thigh length stockings compared with knee length stockings; 54% vs. 20% (Winslow & Brosz 2008). Only 13% of patients wore thigh length stockings satisfactorily compared with 50% of patients wearing knee length

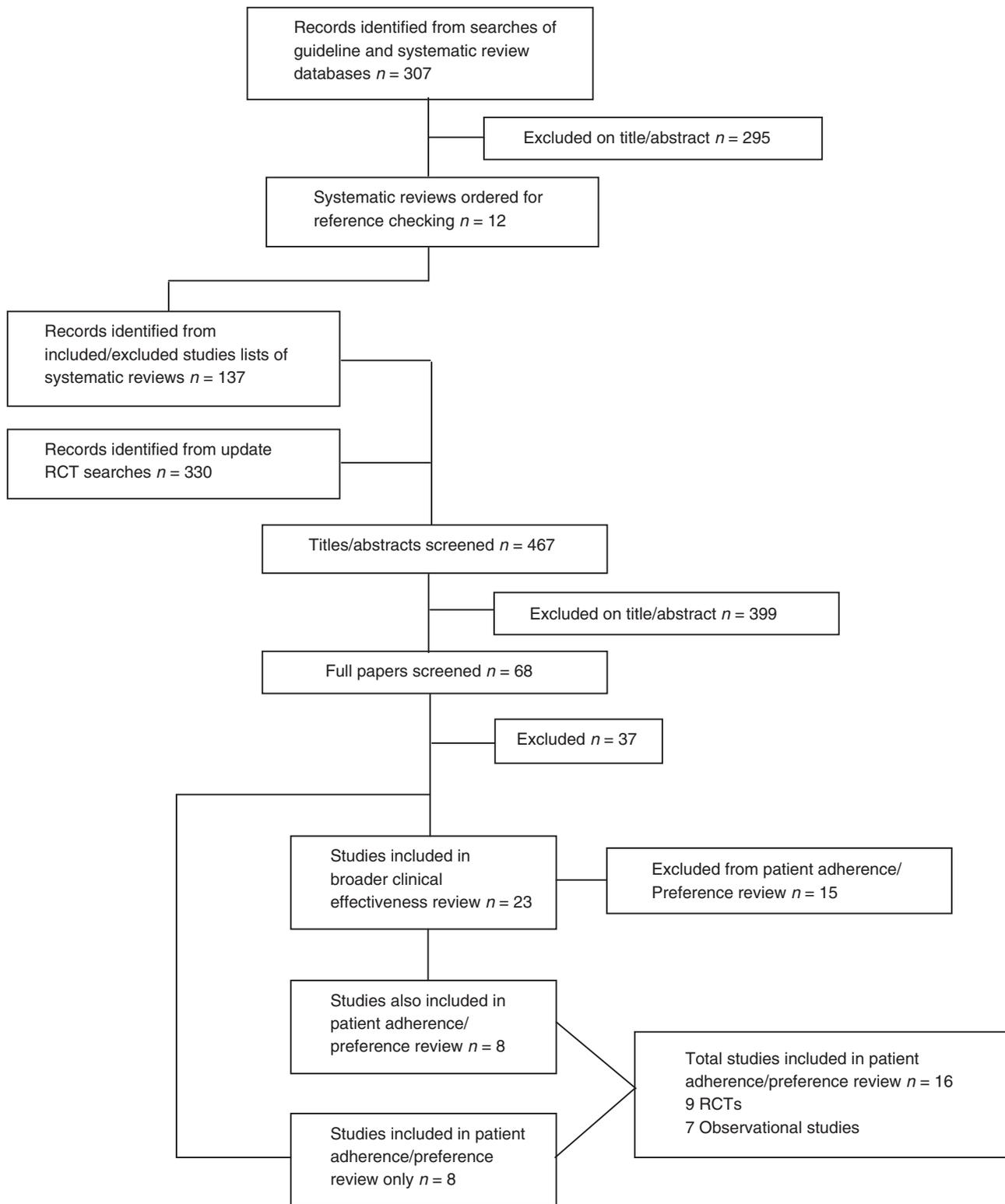
stockings (taking into account sizing, constriction bands and positioning) (Williams *et al.* 1996).

### Patient preference

Three RCTs and five observational studies presented findings on patient preference (Table 4) (Porteous *et al.* 1989, Benko *et al.* 2001, Hameed *et al.* 2002, Parnaby 2004, Williams & Owen 2006, Brady *et al.* 2007, Winslow & Brosz 2008, Ayhan *et al.* 2013). Of the eight studies, six reported on patients' preference for thigh or knee length stockings; the majority of patients in all six studies preferred knee length stockings (Porteous *et al.* 1989, Benko *et al.* 2001, Hameed *et al.* 2002, Williams & Owen 2006, Brady *et al.* 2007, Winslow & Brosz 2008).

One RCT reported that patients found the low pressure knee length stockings 'very comfortable' and the thigh length stockings 'comfortable', but patients reported that moderate pressure knee length stockings were 'uncomfortable' (Ayhan *et al.* 2013). The other two RCTs reported that knee length stockings were more acceptable and more comfortable than thigh length stockings (Porteous *et al.* 1989, Benko *et al.* 2001).

Parnaby (2004) undertook an initial survey of patients wearing a particular brand of knee or thigh length stocking



**Figure 1** Flow diagram of the study selection process.

and two subsequent trials. The second trial administered a modified stocking to incorporate changes to overcome problems identified by patients in the first trial, including a change

in the heel design to prevent excess friction and the availability of an open toe version. Patients preferred the modified stockings and 95% stated that they would wear the stockings

**Table 3** Patient adherence results.

Author	Patient characteristics	Number of patients complying
<b>RCTs</b>		
Benko <i>et al.</i> (2001)	Orthopaedic surgery ( <i>n</i> = 200)	No difference between knee and thigh length stockings for independent management. Thigh length stockings wrinkled significantly more ( $P < 0.05$ ; no confidence intervals presented).
Camporese <i>et al.</i> (2008)	Orthopaedic surgery ( <i>n</i> = 1602)	63 (9.6%) patients in the GCS group, 54 (8.3%) patients in the 7 day LMWH group, and 47 (10.6%) patients in the 14 day LMWH group declined to complete the prophylactic regimen.
Fredin <i>et al.</i> (1989)	Orthopaedic surgery ( <i>n</i> = 144)	2 (4%) patients in the thigh length GCS plus Dextran group discontinued wearing stockings because of discomfort. Control groups received Dextran only.
Hui <i>et al.</i> (1996)	Orthopaedic surgery ( <i>n</i> = 138)	23% of patients in the thigh length stocking group and 16% in the knee length stocking group found the stockings too uncomfortable and requested their removal.
Mellbring and Palmer (1986)	Abdominal surgery ( <i>n</i> = 108)	All patients tolerated wearing thigh length GCS.
Turpie <i>et al.</i> (1989)	Neurosurgery ( <i>n</i> = 239; 173 patients had neurosurgery; 66 patients did not have surgery)	2 (2.5%) patients did not wear thigh length stockings correctly. Ten patients (13%) in the GCS plus intermittent pneumatic compression group did not tolerate treatment although 8 of these patients wore the GCS as required.
Wille-Jorgensen <i>et al.</i> (1985)	Major abdominal surgery ( <i>n</i> = 176)	2 (2.3%) patients removed their thigh length stockings after 5 days, otherwise they were well tolerated.
<b>Observational studies</b>		
Brady <i>et al.</i> (2007)	Mixed surgery ( <i>n</i> = 137)	51 (37%) patients were not wearing stockings (most thigh length). Thirty-four problems with fit on patients wearing thigh length stockings vs 18 problems with knee length stockings.
Hameed <i>et al.</i> (2002)	Mixed surgery ( <i>n</i> = 72)	Higher proportion of patients wearing thigh length stockings incorrectly or with stockings rolled down (64.3%) compared with knee length stockings (30%).
Parnaby (2004)	Mixed surgery: 218 (survey); 70 (trial 1); 20 (trial 2)	119 of 218 (54%) patients were not wearing GCS. One third of patients wearing above knee stockings versus 9% of patients wearing below knee stockings wore them incorrectly.
Thompson <i>et al.</i> (2011)	Orthopaedic surgery <i>n</i> = 56 (audit); <i>n</i> = 57 (trial)	Knee length stockings were incorrectly sized in 28/38 (74%) patients. Twelve of 18 patients (67%) removed knee length stockings as they were uncomfortable or too tight. Implementation of a standardised protocol reduced these problems.
Williams <i>et al.</i> (1994, 1996)	Orthopaedic surgery ( <i>n</i> = 324)	Greater proportion of patients wearing thigh length stockings too low or with wrinkles/bands. Greater satisfaction in patients wearing knee length GCS.
Winslow and Brosz (2008)	Mixed surgery ( <i>n</i> = 142)	Greater proportion of patients in thigh length GCS groups (particularly overweight patients) wearing GCS incorrectly, incorrect size, and reporting skin problems compared with knee length GCS groups.

GCS, graduated compression stockings; LMWH, low molecular weight heparin; RCT, randomised controlled trial.

again. The other four observational studies reported patient preference in terms of comfort, ease of application and general satisfaction (Hameed *et al.* 2002, Williams & Owen 2006, Brady *et al.* 2007, Winslow & Brosz 2008). A greater proportion of patients in the studies preferred knee length stockings, finding them more comfortable.

## Discussion

This review of quantitative evidence on patient preference and adherence to thigh or knee length compression stockings for the prevention of DVT in surgical patients was based on systematic literature searches and combined all

**Table 4** Patient preference results.

Author	Patient characteristics	Patient preference
RCTs		
Ayhan <i>et al.</i> (2013)	Patients at extremely high-risk for postoperative DVT ( $n = 219$ )	Low pressure knee and thigh length stockings reported to be very comfortable/comfortable. Moderate pressure knee length GCS were reported to be uncomfortable ( $P < 0.001$ ; no confidence intervals presented).
Benko <i>et al.</i> (2001)	Orthopaedic surgery ( $n = 200$ )	Significantly more patients reported discomfort with thigh length Thrombex GCS than knee length Thrombex or Brevet TX GCS ( $P < 0.05$ ; no confidence intervals presented).
Porteous <i>et al.</i> (1989)	Major abdominal surgery ( $n = 114$ )	Brevet knee length stockings were more acceptable and comfortable compared to thigh length stockings.
Observational studies		
Brady <i>et al.</i> (2007)	Mixed surgery ( $n = 137$ )	Complaints of discomfort were highest amongst patients wearing thigh length TED stockings compared to knee length TED stockings and/or SCD.
Hameed <i>et al.</i> (2002)	Mixed surgery ( $n = 72$ )	Patients were generally more satisfied with knee length GCS compared to Kendall thigh length GCS.
Parnaby (2004)	Mixed surgery: 218 (survey); 70 (trial 1); 20 (trial 2)	95% patients would wear the modified thigh or knee length SaphenaMedical GCS again. None of the patients wearing the modified heel design complained of heel friction.
Williams and Owen (2006)	Orthopaedic surgery ( $n = 47$ )	All female patients would have preferred to wear knee length compared to thigh length GCS.
Winslow and Brosz (2008)	Mixed surgery ( $n = 142$ )	Higher proportion of patients (including overweight patients) rated thigh length GCS as uncomfortable or very uncomfortable compared to knee length GCS ( $P < 0.001$ ; no confidence intervals presented).

DVT, deep vein thrombosis; GCS, graduated compression stockings; SCD, sequential compression device.

the available RCT evidence, along with the best available observational data on patient preference and adherence to thigh length vs. knee length stockings.

Nine RCTs and seven observational studies were identified that reported data on patient adherence and/or preference with the use of graduated compression stockings post surgery. Patient adherence (wearing stockings correctly and for the required duration) was higher in the RCTs than the observational studies, but across all studies the proportion of patients not wearing stockings or wearing stockings incorrectly (non-adherence) appeared to be generally higher in patients receiving thigh length compared with knee length stockings. All of these studies reflect patient adherence in a hospital setting; it is likely that adherence is even lower after patients have been discharged from hospital. This may have implications on discharge planning to ensure patients are aware of the importance in adhering to wearing stockings correctly and potential use of additional resources to enhance and monitor stocking adherence. In all six studies that reported on patients' preference for length of stocking, patients preferred knee length stockings over thigh length stockings.

Some of the included studies state recommendations for practice and further research, including the three observational studies undertaken by nurses (Parnaby 2004, Brady *et al.* 2007, Winslow & Brosz 2008). Brady states that their results provide additional evidence to the effect that to improve compliance, issues of patient comfort must be considered when creating effective treatment protocols (Brady *et al.* 2007). Parnaby states that the nursing profession should be the main group targeted for improving adherence with graduated compression stocking use, since they are responsible for measuring, fitting and monitoring the patient (Parnaby 2004). Winslow states that nurses should provide patient education when they measure the patient and apply, check and remove stockings; this is very important as many patients continue to wear graduated compression stockings after hospital discharge (Winslow & Brosz 2008). Winslow also makes recommendations for further nursing research, stating that it is important to study patients at home to determine patient compliance after hospital discharge and recommends interviewing nurses who care for patients using compression stockings about several aspects of their practice (Winslow & Brosz 2008).

One of the included RCTs (led by a Research Fellow) and another of the observational studies (led by a medical student) also made recommendations for practice. Benko states that the difficulty orthopaedic patients experience in handling stockings underlines the necessity of regular checks to avoid ischaemic complications, particularly in the high-risk older population (Benko *et al.* 2001). Thompson also emphasises the importance of correct fitting and monitoring of stocking use, stating that a standard protocol of nursing practice is critical to the effectiveness of graduated compression stockings after total hip replacement and total knee replacement (Thompson *et al.* 2011).

### Limitations

This systematic review only included quantitative study designs and did not attempt to identify qualitative research. There was substantial variation between the included studies, in terms of patient characteristics, surgical procedures and methodology. In addition, some studies were very old and may therefore not reflect current practice. Most RCTs had an unclear or high risk of bias and the observational studies were often poorly reported and based on small sample sizes. For some of the results reported in the included studies, *P* values were presented, although estimates of precision (such as confidence intervals) were not reported. Overall, the evidence base of studies assessing surgical patient adherence and preference for thigh length or knee length graduated compression stockings should be considered as weak.

### Conclusion

The evidence on surgical patient adherence and preference for thigh length or knee length graduated compression stockings suggests that patients prefer knee length stockings and are more likely to wear them correctly. However, the evidence base is limited in quality and quantity and lacks 'real life' data on adherence in the community. In many clinical settings any difference in efficacy between thigh length and knee length stockings may be rendered irrelevant by patient preference and likely better adherence with knee length stockings.

The results of our literature review indicate that efforts need to be made to improve patient adherence to the correct use of compression stockings, particularly of thigh length stockings. Nurses have a vital role in patient education to improve patient adherence. It may be helpful for patients to be presented with findings on the effectiveness

of different stocking lengths. They should also be made aware of the related practical issues and any potential adverse effects, such as with thigh length stockings rolling down and constricting blood flow.

Our review also indicates that in practice the choice between thigh and knee length stockings must take into account the likely adherence given each individual patient's particular needs and circumstances. A patient's inability to correctly use thigh length stockings may well outweigh any theoretical efficacy benefit over knee length stockings. The results of our review indicate that any future research into the effectiveness of graduated compression stockings should take into account patient preference and incorporate assessment of patient adherence, both in a hospital setting and post discharge. A synthesis of qualitative evidence would enhance our understanding of patient preferences and reasons for non-adherence to thigh length or knee length compression stockings.

### Acknowledgements

The authors would like to thank Dave Fox, Information Specialist, Centre for Reviews and Dissemination, for conducting the electronic searches. In addition, we would like to thank our patient representative, Mrs Kaye Norman, for advice.

### Funding

This project was commissioned by the NIHR HTA Programme as project number HTA 13/72/01. The views and opinions expressed are those of the authors and do not necessarily reflect those of the NIHR HTA Programme, NIHR, NHS or the Department of Health.

### Conflict of interest

No conflict of interest has been declared by the authors.

### Author contributions

All authors have agreed on the final version and meet at least one of the following criteria [recommended by the ICMJE (<http://www.icmje.org/recommendations/>)]:

- substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data;
- drafting the article or revising it critically for important intellectual content.

## Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site.

## References

- Ayhan H., Iyigun E., Ince S., Can M.F., Hatipoglu S. & Saglam M. (2013) Comparison of three different protocols in the prevention of postoperative deep vein thrombosis in patients at high-risk: randomized clinical study. *European Surgical Research* 50, 64–65.
- Benko T., Cooke E.A., McNally M.A. & Mollan R.A.B. (2001) Graduated compression stockings – knee length or thigh length. *Clinical Orthopaedics and Related Research* 383, 197–203.
- Brady D., Raingruber B., Peterson J., Varnau W., Denman J., Resuello R., De Contreaus R. & Mahnke J. (2007) The use of knee-length versus thigh-length compression stockings and sequential compression devices. *Critical Care Nursing Quarterly* 30, 255–62.
- Camporese G., Bernardi E., Prandoni P., Noventa F., Verlato F., Simioni P., Ntita K., Salmistraro G., Frangos C., Rossi F., Cordova R., Franz F., Zucchetto P., Kontothanassis D. & Andreozzi G.M.; for the KANT (Knee Arthroscopy Nadroparin Thromboprophylaxis) Study Group (2008) Low-molecular-weight heparin versus compression stockings for thromboprophylaxis after knee arthroscopy: a randomized trial. *Annals of Internal Medicine* 149, 73–82.
- Centre for Reviews and Dissemination (2009) *Systematic Reviews: CRD's Guidance for Undertaking Reviews in Health Care*. Centre for Reviews and Dissemination, University of York, York.
- Fredin H., Bergqvist D., Cederholm C., Lindblad B. & Nyman U. (1989) Thromboprophylaxis in hip-arthroplasty – dextran with graded compression or preoperative dextran compared in 150 patients. *Acta Orthopaedica Scandinavica* 60, 678–681.
- Hameed M.F., Browse D.J., Immelman E.J. & Goldberg P.A. (2002) Should knee-length replace thigh-length graduated compression stockings in the prevention of deep-vein thrombosis? *South African Journal of Surgery* 40, 15–16.
- Higgins J. P. T. & Green S. (2011) *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0* [updated March 2011]. The Cochrane Collaboration, Oxford.
- Hui A.C.W., Heraspalou C., Dunn I., Triffitt P.D., Crozier A., Imeson J. & Gregg P.J. (1996) Graded compression stockings for prevention of deep-vein thrombosis after hip and knee replacement. *Journal of Bone and Joint Surgery-British* 78B, 550–554.
- Mellbring G. & Palmer K. (1986) Prophylaxis of deep-vein thrombosis after major abdominal-surgery – comparison between dihydroergotamine-heparin and intermittent pneumatic calf compression and evaluation of added graduated static compression. *Acta Chirurgica Scandinavica* 152, 597–600.
- Moher D., Liberati A., Tetzlaff J., Altman D.G. & Grp P. (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *British Medical Journal* 339, b2535.
- National Institute for Health and Care Excellence (2009) *Medicines Adherence: Involving Patients in Decisions about Prescribed Medicines and Supporting Adherence*. National Institute for Health and Care Excellence, London.
- National Institute for Health and Care Excellence (2010) *NICE Clinical Guidance CG92: Venous Thromboembolism: Reducing the Risk: Reducing the Risk of Venous Thromboembolism (Deep Vein Thrombosis and Pulmonary Embolism) in Patients Admitted to Hospital*. National Institute for Health and Care Excellence, London.
- Nicolaides A.N., Breddin H.K., Fareed J., Goldhaber S., Haas S., Hull R., Kalodiki E., Myers K., Samama M. & Sasahara A. (2001) Prevention of venous thromboembolism. International Consensus Statement. Guidelines compiled in accordance with scientific evidence. *International Angiology* 20, 1–37.
- Parnaby C. (2004) A new anti-embolism stocking. Use of below-knee products and compliance. *British Journal of Perioperative Nursing: The Journal of the National Association of Theatre Nurses* 14, 302–7.
- Porteous M.J.L., Nicholson E.A., Morris L.T., James R. & Negus D. (1989) Thigh length versus knee length stockings in the prevention of deep-vein thrombosis. *British Journal of Surgery* 76, 296–297.
- Sajid M.S., Desai M., Morris R.W. & Hamilton G. (2012) Knee length versus thigh length graduated compression stockings for prevention of deep vein thrombosis in postoperative surgical patients. *Cochrane Database of Systematic Reviews* 5, CD007162. doi: 10.1002/14651858.CD007162.pub2
- Thompson A., Walter S., Brunton L.R., Pickering G.T., Mehendale S., Smith A.J. & Bannister G.C. (2011) Anti-embolism stockings and proximal indentation. *British Journal of Nursing* 20, 1426–1430.
- Turpie A.G.G., Hirsh J., Gent M., Julian D. & Johnson J. (1989) Prevention of deep-vein thrombosis in potential neurosurgical patients - a randomized trial comparing graduated compression stockings alone or graduated compression stockings plus intermittent pneumatic compression with control. *Archives of Internal Medicine* 149, 679–681.
- Wade R., Sideris E., Paton F., Rice S., Palmer S., Fox D., Woolacott N. & Spackman E. (2015) Graduated compression stockings for the prevention of deep vein thrombosis in postoperative surgical patients: a systematic review and economic model with a value of information analysis. *Health Technology Assessment* 19(98), 1–220.
- Wille-Jorgensen P., Thorup J., Fischer A., Holstchristensen J. & Flamsholt R. (1985) Heparin with and without graded compression stockings in the prevention of thromboembolic complications of major abdominal surgery- a randomized trial. *British Journal of Surgery* 72, 579–581.
- Williams L.A. & Owen T.D. (2006) Above-knee versus below-knee stockings in total knee arthroplasty. *Annals of the Royal College of Surgeons of England* 88, 302–305.
- Williams A.M., Rhys-Davies P. & Sweetnam D. (1994) Graduated compression stockings in the prevention of deep vein thrombosis and the case for knee length versus thigh length stockings. *Journal of Bone and Joint Surgery – British* 76B (Suppl. I), 56.

Williams A.M., Davies P.R., Sweetnam D.I.S., Harper G., Pusey R. & Lightowler C.D.R. (1996) Knee-length versus thigh-length graduated compression stockings in the prevention of deep vein thrombosis. *British Journal of Surgery* 83, 1553.

Winslow E.H. & Brosz D.L. (2008) Graduated compression stockings in hospitalized postoperative patients: correctness of usage and size. *American Journal of Nursing* 108, 40–51.

The *Journal of Advanced Nursing (JAN)* is an international, peer-reviewed, scientific journal. *JAN* contributes to the advancement of evidence-based nursing, midwifery and health care by disseminating high quality research and scholarship of contemporary relevance and with potential to advance knowledge for practice, education, management or policy. *JAN* publishes research reviews, original research reports and methodological and theoretical papers.

For further information, please visit *JAN* on the Wiley Online Library website: [www.wileyonlinelibrary.com/journal/jan](http://www.wileyonlinelibrary.com/journal/jan)

**Reasons to publish your work in *JAN*:**

- **High-impact forum:** the world's most cited nursing journal, with an Impact Factor of 1.917 – ranked 8/114 in the 2015 ISI Journal Citation Reports © (Nursing (Social Science)).
- **Most read nursing journal in the world:** over 3 million articles downloaded online per year and accessible in over 10,000 libraries worldwide (including over 3,500 in developing countries with free or low cost access).
- **Fast and easy online submission:** online submission at <http://mc.manuscriptcentral.com/jan>.
- **Positive publishing experience:** rapid double-blind peer review with constructive feedback.
- **Rapid online publication in five weeks:** average time from final manuscript arriving in production to online publication.
- **Online Open:** the option to pay to make your article freely and openly accessible to non-subscribers upon publication on Wiley Online Library, as well as the option to deposit the article in your own or your funding agency's preferred archive (e.g. PubMed).

# Interventional management of hyperhidrosis in secondary care: a systematic review

R. Wade<sup>1</sup>, A. Llewellyn,<sup>1</sup> J. Jones-Diette,<sup>1</sup> K. Wright,<sup>1</sup> S. Rice,<sup>2</sup> A.M. Layton,<sup>3</sup> N.J. Levell,<sup>4</sup> D. Craig<sup>2</sup> and N. Woolcott<sup>1</sup>

<sup>1</sup>Centre for Reviews and Dissemination, University of York, Heslington, York YO10 5DD, U.K.

<sup>2</sup>Institute of Health and Society, Newcastle University, Newcastle upon Tyne, U.K.

<sup>3</sup>Harrogate and District NHS Foundation Trust, Harrogate, U.K.

<sup>4</sup>Norfolk and Norwich University Hospitals NHS Foundation Trust, Norwich, U.K.

## Summary

### Correspondence

Ros Wade.

E-mail: ros.wade@york.ac.uk

### Accepted for publication

6 March 2018

### Funding sources

This project was funded by the National Institute for Health Research HTA Programme (project number 14/211/02). The views and opinions expressed herein are those of the authors and do not necessarily reflect those of the Department of Health.

### Conflicts of interest

None to declare.

DOI 10.1111/bjd.16558

**Background** Hyperhidrosis is uncontrollable excessive sweating, which occurs at rest, regardless of temperature. The symptoms of hyperhidrosis can significantly affect quality of life.

**Objectives** To undertake a systematic review of the clinical effectiveness and safety of treatments available in secondary care for the management of primary hyperhidrosis.

**Methods** Fifteen databases (including trial registers) were searched to July 2016 to identify studies of secondary-care treatments for primary hyperhidrosis. For each intervention randomized controlled trials (RCTs) were included where available; where RCT evidence was lacking, nonrandomized trials or large prospective case series were included. Outcomes of interest included disease severity, sweat rate, quality of life, patient satisfaction and adverse events. Trial quality was assessed using a modified version of the Cochrane Risk of Bias tool. Results were pooled in pairwise meta-analyses where appropriate, otherwise a narrative synthesis was presented.

**Results** Fifty studies were included in the review: 32 RCTs, 17 nonrandomized trials and one case series. The studies varied in terms of population, intervention and methods of outcome assessment. Most studies were small, at high risk of bias and poorly reported. The interventions assessed were iontophoresis, botulinum toxin (BTX) injections, anticholinergic medications, curettage and newer energy-based technologies that damage the sweat gland.

**Conclusions** The evidence for the effectiveness and safety of treatments for primary hyperhidrosis is limited overall, and few firm conclusions can be drawn. However, there is moderate-quality evidence to support the use of BTX for axillary hyperhidrosis. A trial comparing BTX with iontophoresis for palmar hyperhidrosis is warranted.

### What's already known about this topic?

- Hyperhidrosis is characterized by uncontrollable excessive sweating, which occurs at rest, regardless of temperature; symptoms can significantly affect quality of life.
- Hyperhidrosis with no discernible cause is known as primary hyperhidrosis.
- Despite the existence of a wide range of treatments for primary hyperhidrosis and a large number of clinical studies, there is uncertainty regarding optimal patient management and substantial variation in the availability of secondary-care treatments in the U.K.

### What does this study add?

- This high-quality systematic review synthesizes the large amount of research evidence for the effectiveness and safety of treatments for primary hyperhidrosis, which unfortunately is of limited quality, and few firm conclusions can be drawn.
- There is moderate-quality evidence to support the use of botulinum toxin injections for axillary hyperhidrosis.
- Recommendations for robust research are made, based on the results of the systematic review, alongside clinical and patient advice.

Hyperhidrosis is characterized by uncontrollable excessive and unpredictable sweating, which occurs at rest, regardless of temperature. Primary hyperhidrosis, which is the focus of this review, has no discernible cause. It most commonly involves the axillae, palms and soles, but may also involve the face, groin or any area of the body.

Primary hyperhidrosis is thought to affect at least 1% of the U.K. population.<sup>1</sup> The symptoms of hyperhidrosis can significantly affect quality of life, and can lead to social embarrassment, loneliness, anxiety and depression. It can impair work activities or studying in those handling pens, paper and electronic equipment. Functional problems may arise from skin maceration and soreness. Severely affected patients may also have secondary microbial infections. The unpredictable and uncontrollable nature of the condition can make it very distressing for patients.

In primary care, patients may initially be advised to make lifestyle changes such as restricting stimulant-containing foods, losing weight and avoiding clothing that can make sweating worse. First-line treatment includes topical pharmacological agents: aluminium chloride has been shown to be effective for mild-to-moderate axillary hyperhidrosis and formaldehyde solution can be prescribed for plantar hyperhidrosis.<sup>2,3</sup> Unfortunately, skin irritation is very common with these antiperspirants and often forces discontinuation of the treatment.<sup>4</sup>

Patients may be referred to a dermatologist if treatment fails or is not tolerated. However, current recommendations are not underpinned by robust evidence and there is significant variation in the availability of treatments for primary hyperhidrosis in secondary care in the U.K. Further clinical trials may be required, in particular comparing the effectiveness of treatments prescribed by a dermatologist, but first a thorough review of the available evidence is warranted.

The aim of this study was to undertake a systematic review of the clinical effectiveness and safety of treatments available in secondary care for the management of patients with refractory primary hyperhidrosis.

## Methods

A protocol for the systematic review was developed and registered on PROSPERO (number CRD42015027803). The review

included studies of patients (adults and children) with primary hyperhidrosis. Studies of any treatment for hyperhidrosis offered in secondary care for prescription by dermatologists, and minor surgical treatments, were eligible for inclusion. Endoscopic thoracic sympathectomy was not included as it is not recommended by many practitioners; it is generally considered only as an intervention of last resort due to its significant risks and common adverse effects such as compensatory hyperhidrosis.<sup>5</sup>

For each intervention randomized controlled trials (RCTs) were included, where available. For interventions where RCT evidence was lacking, non-RCTs or large prospective case series were included. Recently published high-quality systematic reviews were also considered if they were directly relevant. Outcomes of interest included disease severity, sweat rate, quality of life, patient satisfaction and adverse events.

Potentially relevant studies were identified through literature searching. Twelve databases were searched in January 2016 (including MEDLINE, Embase and the Cochrane Central Register of Controlled Trials). No date or language limits were applied. The MEDLINE search strategy, which identified the greatest number of records, is presented in Appendix S1 (see Supporting Information). Clinical advisors were consulted for additional studies, and reference lists of relevant systematic reviews were manually searched. Information on studies in progress and unpublished research was sought by searching conference proceedings and trial registers, in July 2016.

Two researchers (R.W. and J.J.-D.) undertook the screening of titles and abstracts obtained through the search, although the library was split between the researchers, rather than each record being double screened. A sample of just over 10% of records was double screened in order to assess the level of agreement between the researchers; it was planned to undertake full double screening if the level of agreement was poor, but this was not necessary as the level of agreement between researchers was 96.2%. Full manuscripts of potentially relevant studies were obtained and independently screened by two researchers (R.W. and J.J.-D.), using predefined eligibility criteria. Disagreements were resolved through discussion or consultation with a third researcher. Relevant foreign-language studies were translated and included.

Data were extracted directly into a standardized, piloted spreadsheet developed in Microsoft Excel (by R.W., A.L. and

J.J.-D.). Data extracted included study design, sample size, participant characteristics (body site treated, age, sex, previous treatments, baseline disease severity), treatment characteristics (dose, frequency, duration), outcomes assessed (measurement tool and time point) and results. Data extraction was conducted by one researcher and checked for accuracy by a second. In cases of multiple publications of the same study, the publication with the largest sample or longest follow-up was treated as the main source. Where possible we extracted intention-to-treat data. Where results data were missing or limited (e.g. only presented in graphical format, or conference abstracts), authors were contacted and, where relevant, manufacturer trials registers were consulted for further data. If the authors did not respond, data from graphs were extracted using Graph Grabber software (Quintessa, Henley-on-Thames, U.K.).

The quality of RCTs and non-RCTs was assessed using a modified version of the Cochrane Risk of Bias tool by one researcher and checked for accuracy by a second (R.W., A.L. and J.J.-D.).<sup>6</sup> An additional question relating to the similarity of treatment groups at baseline was added.<sup>7</sup> In addition, a question about 'within-patient' study designs was added, owing to concerns about the validity of certain outcome measures in 'within-patient' study designs, in which patients receive different interventions on different sides of the body (i.e. the left vs. right axilla). The results of the risk-of-bias assessment are shown in Appendix S2 (see Supporting Information). Case series were not formally quality assessed; their results were presented as supporting evidence. No systematic reviews were included in the review except as a source of relevant studies, so they were not quality assessed.

Results were pooled in pairwise meta-analyses if at least two studies of the same intervention and comparator reported the same outcome and were considered sufficiently similar for analysis to be appropriate and feasible. Otherwise, results were summarized in a narrative synthesis. Where meta-analyses were performed, dichotomous outcomes were combined to estimate pooled risk ratios (RRs), and continuous outcomes were combined to estimate pooled mean differences (MDs) using random effects DerSimonian–Laird meta-analyses.<sup>8</sup> Statistical heterogeneity was assessed using the  $I^2$ -statistic and visual inspection of forest plots. Studies using different units of analysis (i.e. axilla in half-side comparisons vs. patients in between-patient comparisons) were pooled where deemed appropriate and reported in separate subgroups.

For studies that included two separate intervention groups with two different doses and used one control group, data from each intervention group were entered separately to explore any dose–response effect, and the number of participants in the control group was divided by two to reduce the risk of double counting data.<sup>9</sup> Although this approach may artificially reduce the power of the study in the meta-analysis and does not account for potential correlation between the two active treatment groups, a separate analysis combining the two arms showed no significant difference in results.

Metaregressions and other subgroup analyses were considered inappropriate due to the small number of studies. All

analyses were conducted using the Cochrane Collaboration's Review Manager 5.3. Clinical and patient advisors contributed to the interpretation of the results.

## Results

The electronic searches identified a total of 4057 records; the flow diagram of the study selection process is presented in Figure 1.

Appendix S3 (see Supporting Information) presents the 155 records that met the inclusion criteria for the systematic review. For each intervention for which there were RCTs or nonrandomized comparative studies available, less robust studies were excluded, resulting in 93 small case series being excluded from the review. Five additional studies were excluded because they were systematic reviews that were not considered to be of sufficiently good quality, up to date or directly relevant enough to be relied upon, resulting in 57 records (reporting 48 studies) identified for inclusion in the review.

An additional two studies were identified from the separate searches of conference proceedings and trial registers (flow diagram presented in Appendix S4; see Supporting Information). Therefore, in total 50 studies were included in the review: 32 RCTs, 17 non-RCTs and one case series.

## Study characteristics

The studies varied in terms of country of origin (indicating climate and population differences), intervention and the methods of outcome assessment. Most studies were small (sample sizes ranged from four to 339, with most studies including fewer than 50 patients), at high risk of bias and poorly reported. Further details are provided in Appendix S2 (see Supporting Information). The interventions assessed were iontophoresis, botulinum toxin (BTX), anticholinergic medications, curettage and newer technologies that damage the sweat gland. The majority of studies included only adult patients, and the majority of participants across the studies were female. Where reported, baseline disease severity was moderate to severe, with a Hyperhidrosis Disease Severity Scale (HDSS) score of 3–4 and/or a sweat rate of  $\geq 50$  mg per 5 min. The site of hyperhidrosis differed between studies of different interventions. A summary of the study characteristics is presented in Table S1 (see Supporting Information), with further details presented in Appendix S5 (see Supporting Information).

## Clinical effectiveness

This section presents a summary of the results, presented by intervention. Further results of each study are presented in Appendix S5 (see Supporting Information).

## Iontophoresis

Ten studies (four RCTs, five non-RCTs, one case series) of iontophoresis were included.<sup>10–19</sup> All were at a high or

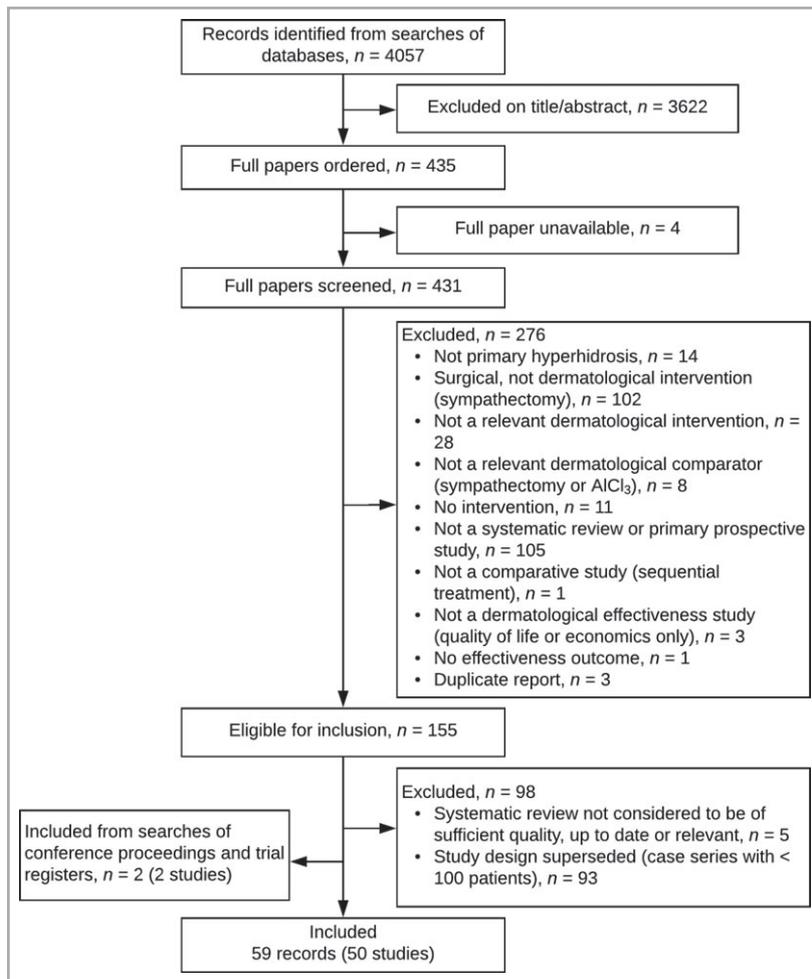


Fig 1. Flow diagram of the study selection process.

unclear risk of bias. There were a number of differences in the iontophoresis interventions used across these studies, with variations in the medium used (tap water, with aluminium chloride or an anticholinergic added, or a 'dry type' device), the electric current used, and the frequency of iontophoresis sessions. No meta-analysis was possible owing to the differences between interventions and outcomes assessed.

Three very small studies (two RCTs and one interrupted time series) with short follow-up times compared tap-water iontophoresis with placebo for palmar hyperhidrosis<sup>10–12</sup> and found a positive effect of iontophoresis as assessed by gravimetry or iodine starch test. This finding was supported by a larger case series.<sup>13</sup>

Of two small nonrandomized comparisons of a handheld 'dry type' iontophoresis device compared with no treatment,<sup>14,15</sup> only one found a statistically significant reduction in sweating, assessed by gravimetry.<sup>14</sup>

Two studies compared iontophoresis alone with iontophoresis combined with anticholinergic therapy for palmo-plantar hyperhidrosis; one RCT found no significant benefit with the addition of oral oxybutynin,<sup>16</sup> while a non-RCT reported that iontophoresis with topical glycopyrrolate

resulted in a longer duration of effect.<sup>17</sup> The addition of anticholinergic therapy was associated with dry throat, mouth or eyes in some patients.

Two studies (one RCT, one non-RCT)<sup>18,19</sup> compared iontophoresis with BTX injections for palmar hyperhidrosis. The RCT found a statistically and clinically significant difference in treatment response (HDSS) and patient-reported symptoms between the two interventions, favouring BTX at 4 weeks from baseline.<sup>18</sup> This result was supported by the non-RCT, but the difference in treatment benefit was no longer statistically significant at 6 or 12 months.<sup>19</sup> Patients receiving BTX were more likely to report mild-to-moderate pain associated with treatment.

Overall, there is very low-quality but consistent evidence suggesting a short-term beneficial effect of tap-water iontophoresis in the treatment of palmar hyperhidrosis. There is inconsistent evidence regarding the beneficial effect of adding anticholinergic therapy to iontophoresis for palmo-plantar hyperhidrosis. There is very low-quality evidence suggesting that BTX is more effective than iontophoresis for palmar hyperhidrosis in the short term. No serious adverse events related to iontophoresis were reported.

**Botulinum toxin (subcutaneous injection)**

Twenty-three studies of BTX, delivered by subcutaneous injection, were included. There was some variation in the BTX used in these trials. Most studies used BTX type A, and only two used type B. Where stated, the most common dose of BTX-A was 50 U, although some studies used up to 250 U. The studies of BTX-B used 2500 U or 5000 U.

For axillary hyperhidrosis, BTX was compared with placebo in nine studies (eight RCTs,<sup>20–27</sup> one open-label continuation study),<sup>28</sup> no treatment in three studies (non-RCTs)<sup>29–31</sup> and curettage in four studies (one RCT,<sup>32</sup> three non-RCTs).<sup>33–35</sup>

For the comparison with placebo, meta-analysis of some trials was possible for the following outcomes: patient-reported symptom improvement [HDSS reduction of at least 2 points, RR 3.30, 95% confidence interval (CI) 2.46–4.43,  $P < 0.001$ ,  $I^2 = 0\%$ ; two studies] (Fig. 2); sweat reduction (gravimetry) expressed as MDs (MD at 16 weeks:  $-66.9$ , 95% CI  $-82.8$  to  $-51.1$ ,  $P < 0.001$ ,  $I^2 = 0\%$ , three studies) (Figs 3, 4) or RRs (RR at 16 weeks: 2.87, 95% CI 1.94–4.26,  $P < 0.001$ ,  $I^2 = 48\%$ , three studies) (Figs 5–7); and quality of life (MD  $-4.80$ , 95% CI  $-5.67$  to  $-3.94$ ,  $P < 0.001$ ,  $I^2 = 3\%$ ; two studies) (Fig. 8).

Overall, the meta-analyses showed a large and clinically significant effect of BTX for axillary hyperhidrosis; benefits were largely sustained at 16 weeks of follow-up (Figs 4, 6). The placebo-controlled BTX trials that were not included in the meta-analyses also reported clinically relevant improvements in sweating<sup>26,27</sup> and improvements in quality of life.<sup>21,28,36</sup> No serious or severe treatment-related adverse events were reported; the most common treatment-related adverse events were injection-site pain and compensatory sweating.

The three non-RCTs comparing BTX with no treatment reported broadly similar results: significant reductions in sweating but injection-site pain associated with BTX injections.<sup>29–31</sup>

The results of the studies comparing BTX with curettage are described in the ‘Curettage’ section below.

For palmar hyperhidrosis, BTX was compared with placebo in three RCTs, which reported a small statistically significant reduction in sweating at 3–13 weeks, measured by gravimetry<sup>37</sup> or sweat area,<sup>38</sup> but not by iodine starch test.<sup>36</sup> Patients’ assessment of disease severity was statistically significantly improved

in the BTX group in all three RCTs. One of the RCTs reported a high incidence of treatment-related adverse events, including decreased grip strength, muscle weakness and dry mouth.<sup>36</sup> Two nonrandomized studies compared BTX with no treatment,<sup>30,39</sup> the results were similar to the findings of the RCTs.

Overall, there is moderate-quality evidence of a large statistically significant effect of BTX injections on symptoms of axillary hyperhidrosis in the short and medium term (up to 16 weeks) compared with placebo. Short-term evidence indicated that BTX may improve quality of life compared with placebo. BTX is associated with mild adverse events, notably injection-site pain. Evidence comparing the effectiveness of BTX injections to the axillae with curettage is very low quality and uncertain. There is very low-quality evidence suggesting that BTX injections had a small positive effect on palmar hyperhidrosis symptoms compared with placebo or no treatment, although adverse events were reported. As stated above, there is very low-quality evidence suggesting that BTX is more effective than iontophoresis for palmar hyperhidrosis in the short term. There is insufficient evidence on the effect of BTX injections on quality of life in palmar hyperhidrosis.

**Topical botulinum toxin**

Only one very small placebo-controlled RCT (unclear risk of bias) evaluated the efficacy of topically applied BTX for axillary hyperhidrosis; there was a greater reduction in sweating with BTX than with placebo.<sup>40</sup> Therefore there is insufficient evidence to conclude on the effectiveness and safety of topical BTX for primary hyperhidrosis.

**Anticholinergics**

Studies of three anticholinergics were identified: topical glycopyrrolate, oral oxybutynin and oral methantheline bromide. No meta-analysis was possible owing to the differences between interventions and outcomes assessed. Two small low-quality RCTs (with high or unclear risk of bias) evaluated short-term treatment with glycopyrrolate wipes against placebo, used for hyperhidrosis of the axilla<sup>41</sup> or the face.<sup>42</sup> Both studies found a significant treatment benefit in terms of

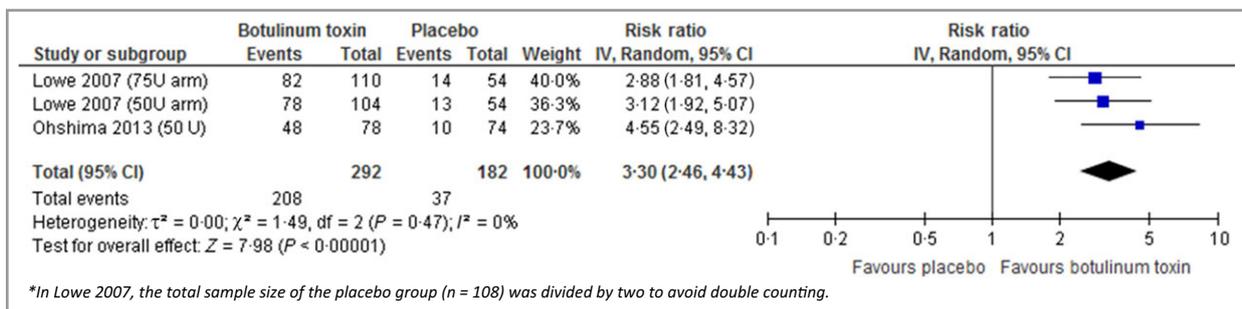


Fig 2. Botulinum toxin vs. placebo: reduction of  $\geq 2$  points in Hyperhidrosis Disease Severity Scale at 4 weeks. CI, confidence interval.

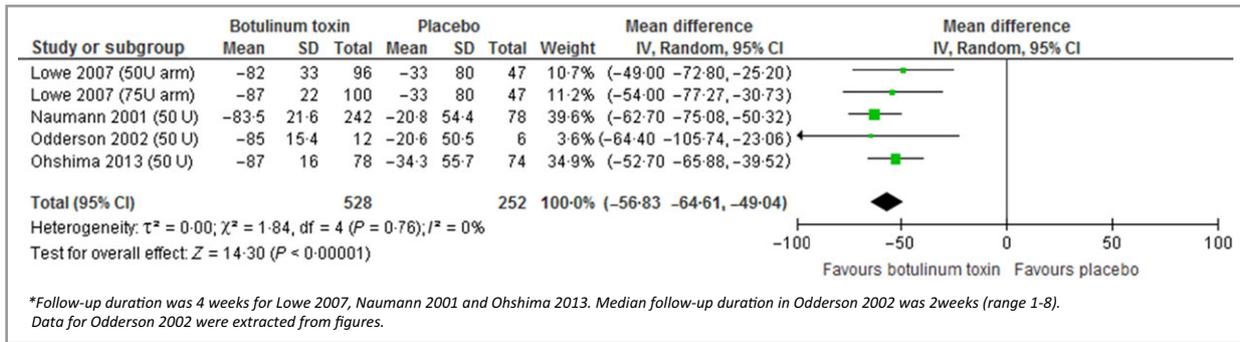


Fig 3. Botulinum toxin vs. placebo: mean percentage change from baseline in sweating at 2–4 weeks. CI, confidence interval.

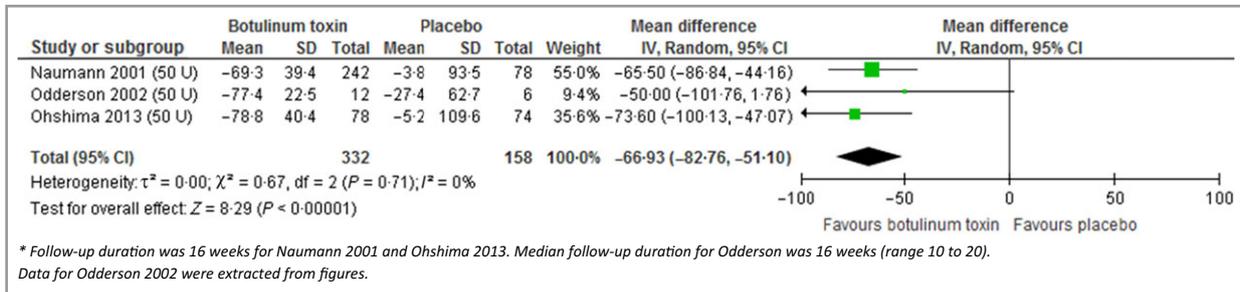


Fig 4. Botulinum toxin vs. placebo: mean percentage change from baseline in sweating at 16 weeks. CI, confidence interval.

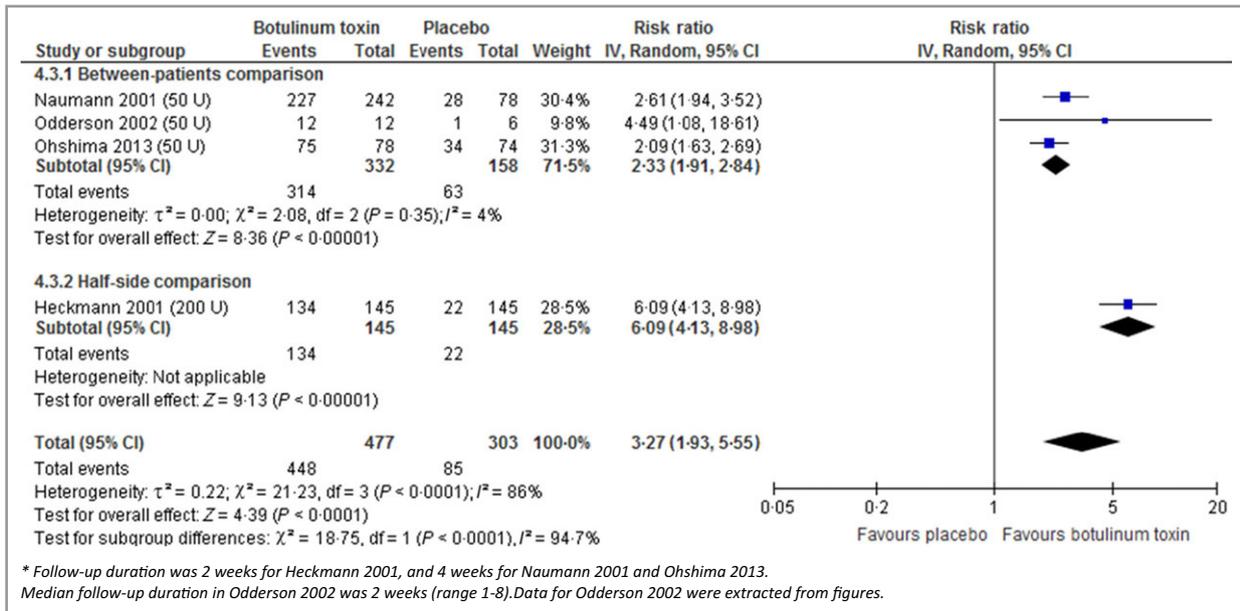


Fig 5. Botulinum toxin vs. placebo: reduction of  $\geq 50\%$  sweating from baseline at 2–4 weeks. CI, confidence interval.

sweating (gravimetry), but improvement in HDSS was seen only in patients receiving treatment for axillary hyperhidrosis.<sup>41</sup> There was limited and inconclusive evidence from one non-RCT<sup>43</sup> regarding the effectiveness (HDSS) and safety of glycopyrrolate spray compared with BTX injections for axillary

hyperhidrosis. There were no studies assessing the clinical effectiveness of oral glycopyrrolate.

Three placebo-controlled RCTs evaluated the effectiveness and safety of oral oxybutynin for hyperhidrosis of the axilla and palm<sup>44</sup> or foot,<sup>45</sup> and generalized hyperhidrosis,<sup>46</sup> and

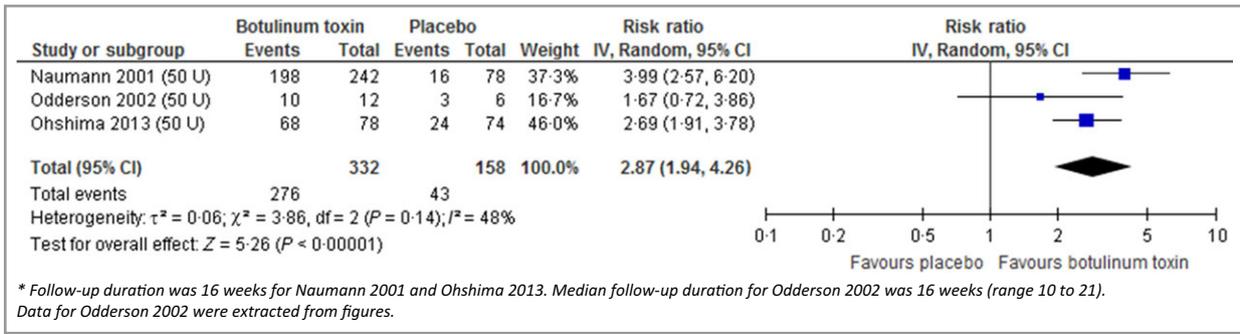


Fig 6. Botulinum toxin vs. placebo: reduction of  $\geq 50\%$  sweating from baseline at 16 weeks. CI, confidence interval.

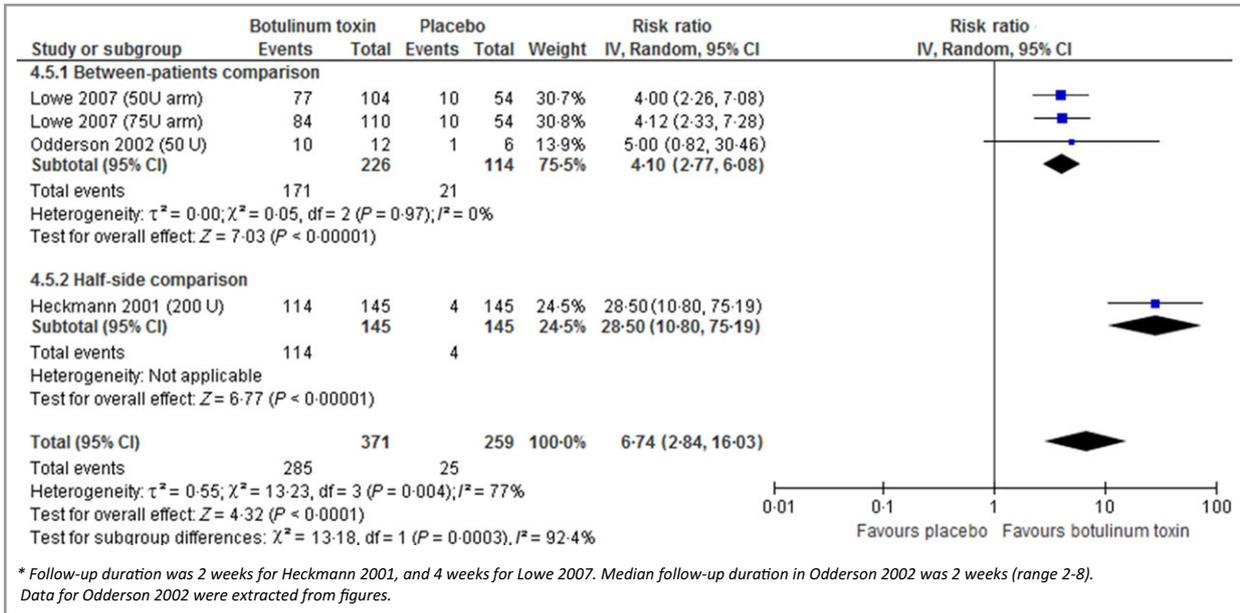


Fig 7. Botulinum toxin vs. placebo: reduction of  $\geq 75\%$  sweating from baseline at 2–4 weeks. CI, confidence interval.

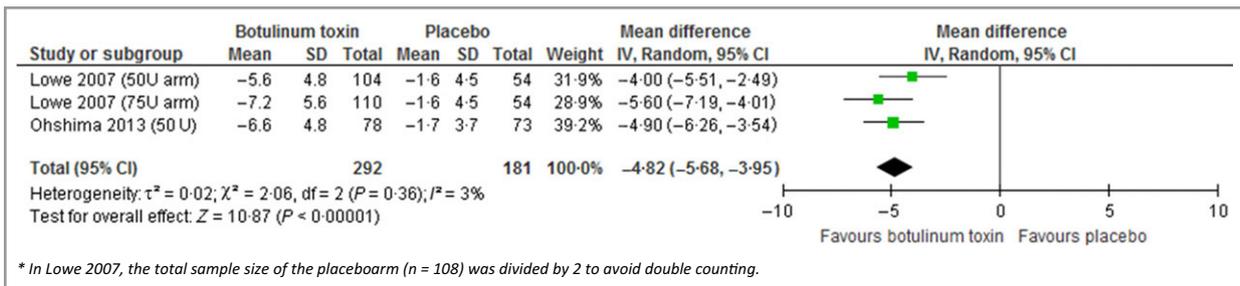


Fig 8. Botulinum toxin vs. placebo: mean change from baseline in Dermatology Life Quality Index score at 4 weeks. CI, confidence interval.

two placebo-controlled RCTs assessed oral methantheline bromide for axillary and palmar hyperhidrosis.<sup>47,48</sup> All studies were at high or unclear risk of bias and reported treatment benefits, as well as a significantly higher incidence of dry-mouth symptoms in patients receiving active therapy.

Overall, the evidence for anticholinergic medications was limited, but suggested short-term benefits of topical glycopyrrolate, oral oxybutynin and oral methantheline bromide on hyperhidrosis symptoms. Oral oxybutynin and methantheline bromide were also associated with dry-mouth adverse events.

## Curettage

Nine studies (four RCTs, five non-RCTs) evaluated curettage for axillary hyperhidrosis. All were at high risk of bias. No meta-analysis was possible owing to the differences between interventions and outcomes assessed.

Of the four studies (one RCT, three non-RCTs) that compared curettage with BTX in axillary hyperhidrosis,<sup>32–35</sup> only the small RCT<sup>32</sup> found a statistically significant difference in HDSS score (at 3 and 6 months of follow-up) favouring BTX. The other studies found no significant difference between treatment groups in sweating, quality-of-life and satisfaction outcomes. However, where reported, the incidence of adverse events was higher with curettage than with BTX.

Five studies (three RCTs, two non-RCTs) compared suction curettage with other surgical interventions: radical skin excision; liposuction curettage, radical skin excision and a skin-sparing technique (Shelley radical skin excision); curettage with and without aggressive manual shaving; tumescent suction curettage and laser.<sup>49–53</sup> Overall, there is very low-quality evidence regarding the relative effectiveness and safety of curettage compared with other minor surgical interventions for axillary hyperhidrosis. Compared with the more radical excision techniques, there is insufficient evidence to demonstrate a clinically significant difference in sweat reduction, patient satisfaction or safety.

## Energy-based ‘destructive’ technologies

Three RCTs evaluated the efficacy and safety of laser epilation for axillary hyperhidrosis.<sup>53–55</sup> All were at high risk of bias and, as well as other study differences, the wavelength used varied between the studies. No meta-analysis was possible owing to the differences between interventions and outcomes assessed. One RCT compared laser with curettage (described in the ‘Curettage’ section above).<sup>53</sup> Two small RCTs compared laser epilation with no treatment; one found that sweating was visibly reduced on the laser-treated side compared with the untreated side at 1 month,<sup>55</sup> but the other study found no significant difference between the treated and untreated sides in sweat reduction at 12 months.<sup>54</sup> Both studies reported no serious adverse events.

One nonrandomized study (high risk of bias) compared the efficacy of fractionated microneedle radiofrequency with a sham control for axillary hyperhidrosis.<sup>56</sup> The study reported significantly better results in mean HDSS scores and sweating intensity at the 21-week follow-up, with transient but not severe adverse events.

One RCT (high risk of bias) compared a microwave device with sham treatment for axillary hyperhidrosis.<sup>57</sup> The study found that microwave therapy was more effective than placebo at reducing patient-reported disease severity, although there was no evidence of a significant difference in the proportion of patients achieving 50% sweat reduction at up to 6 months. Adverse events were generally transient and none was considered severe.

Two small RCTs (high risk of bias) compared microfocused ultrasound with sham treatment for axillary hyperhidrosis, reported in a single publication.<sup>58</sup> The studies reported some benefit in terms of sweating and HDSS.

Overall, there is insufficient evidence regarding the safety and effectiveness of laser epilation, fractionated microneedle radiofrequency, microwave therapy or ultrasound therapy for axillary hyperhidrosis.

## Discussion

The evidence for the effectiveness and safety of second-line treatments for primary hyperhidrosis is limited overall. Most of the included studies were small, at high risk of bias and poorly reported; only one RCT was judged to have a low overall risk of bias. There was insufficient evidence to draw firm conclusions regarding the relative effectiveness and safety of most of the available treatments for primary hyperhidrosis in secondary care.

However, there is moderate-quality evidence of a large effect of BTX injections on symptoms of axillary hyperhidrosis in the short to medium term, although injections were associated with transient injection-site pain. Evidence for other interventions is of low or very low quality. Although the evidence for iontophoresis is very low quality, it is consistent, suggesting that there is a short-term beneficial effect of tap-water iontophoresis in the treatment of palmar hyperhidrosis; no serious adverse events were reported.

There is very low-quality evidence suggesting short-term benefits of topical glycopyrrolate, oral oxybutynin and oral methantheline bromide on hyperhidrosis symptoms. However, oral oxybutynin and methantheline bromide were associated with dry-mouth adverse events. There were no studies assessing the clinical effectiveness of oral glycopyrrolate or propantheline bromide for hyperhidrosis, despite these being commonly used anticholinergic drugs in hyperhidrosis. There was insufficient evidence to demonstrate a clinically significant difference between curettage and other minor surgical interventions or BTX for axillary hyperhidrosis. Evidence was very limited regarding the newer energy-based ‘destructive’ technologies.

Despite its large volume the poor quality of much of the available research evidence is a limitation of this review. The only comparison for which adequate data were available to undertake meta-analysis was that between BTX and placebo for axillary hyperhidrosis. It was not feasible to undertake network meta-analysis; therefore, the comparative clinical effectiveness of the available treatments could not be estimated. In addition, the substantial variation among the included studies limits the generalizability and reliability of the results.

There is limited but promising evidence for the effectiveness of BTX for palmar hyperhidrosis, and therefore a well-conducted, adequately powered RCT of BTX (with anaesthesia) compared with iontophoresis (as the current standard treatment for palmar hyperhidrosis in many dermatology units) for palmar hyperhidrosis may be warranted. This trial should evaluate patient-relevant outcomes based on a validated

scale such as the new HidroQoL© tool. The cost of BTX plus anaesthesia is considerably higher than that of iontophoresis; therefore, the relative cost-effectiveness of these treatments should also be assessed.

In conclusion, the evidence for the effectiveness and safety of treatments for primary hyperhidrosis is limited overall, and few firm conclusions can be drawn. However, there is moderate-quality evidence to support the use of BTX injections for axillary hyperhidrosis. A trial comparing BTX injections with iontophoresis for palmar hyperhidrosis is warranted.

## Acknowledgments

We would like to thank Mr Gerard Stansby, Consultant Vascular Surgeon at the Newcastle upon Tyne Hospitals NHS Foundation Trust; Ms Julie Halford, Specialist Nurse at the Hampshire Clinic for clinical advice; and Dr Mark Simmonds, Research Fellow, Centre for Reviews and Dissemination for statistical advice. We would also like to thank the four patient advisors for their comments and advice.

## References

- Hyperhidrosis Support Group. What is hyperhidrosis? Available at: <https://hyperhidrosisuk.org/what-is-hyperhidrosis> (last accessed 19 June 2018).
- National Institute for Health and Care Excellence. Clinical knowledge summary on hyperhidrosis. Available at: <http://cks.nice.org.uk/hyperhidrosis> (last accessed 19 June 2018).
- Benson RA, Palin R, Holt PJ *et al.* Diagnosis and management of hyperhidrosis. *BMJ* 2013; **347**:f6800.
- Lee KY, Levell NJ. Turning the tide: a history and review of hyperhidrosis treatment. *JRSM Open* 2014; **5**:2042533313505511.
- Smidfelt K, Drott C. Late results of endoscopic thoracic sympathectomy for hyperhidrosis and facial blushing. *Br J Surg* 2011; **98**:1719–24.
- Higgins J, Altman D, Gotzsche P *et al.* The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ Clin Res Ed* 2011; **343**:5928.
- Corbett M, Higgins J, Woolcott N. Assessing baseline imbalance in randomised trials: implications for the Cochrane risk of bias tool. *Res Synth Methods* 2014; **5**:79–85.
- DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986; **7**:177–88.
- Higgins J, Green S. *Cochrane Handbook for Systematic Reviews of Interventions*. Available at: [www.handbook.cochrane.org](http://www.handbook.cochrane.org) (last accessed 19 June 2018).
- Dahl JC, Glent-Madsen L. Treatment of hyperhidrosis manuum by tap water iontophoresis. *Acta Derm Venereol* 1989; **69**:346–8.
- Karakoç Y, Aydemir EH, Kalkan MT. Placebo-controlled evaluation of direct electrical current administration for palmoplantar hyperhidrosis. *Int J Dermatol* 2004; **43**:503–5.
- Stolman LP. Treatment of excess sweating of the palms by iontophoresis. *Arch Dermatol* 1987; **123**:893–6.
- Karakoç Y, Aydemir EH, Kalkan MT, Unal G. Safe control of palmoplantar hyperhidrosis with direct electrical current. *Int J Dermatol* 2002; **41**:602–5.
- Na GY, Park BC, Lee WJ *et al.* Control of palmar hyperhidrosis with a new 'dry-type' iontophoretic device. *Dermatol Surg* 2007; **33**:57–61.
- Choi YH, Lee SJ, Kim DW *et al.* Open clinical trial for evaluation of efficacy and safety of a portable 'dry-type' iontophoretic device in treatment of palmar hyperhidrosis. *Dermatol Surg* 2013; **39**:578–83.
- Shimizu H, Tamada Y, Shimizu J *et al.* Effectiveness of iontophoresis with alternating current (AC) in the treatment of patients with palmoplantar hyperhidrosis. *J Dermatol* 2003; **30**:444–9.
- Dolianitis C, Scarff CE, Kelly J *et al.* Iontophoresis with glycopyrrrolate for the treatment of palmoplantar hyperhidrosis. *Australas J Dermatol* 2004; **45**:208–12.
- Rajagopal R, Mallya NB. Comparative evaluation of botulinum toxin versus iontophoresis with topical aluminium chloride hexahydrate in treatment of palmar hyperhidrosis. *Med J Armed Forces India* 2014; **70**:247–52.
- Wachal K, Bucko W, Staniszewski R *et al.* Estimating the subjective efficiency after treatment of upper extremity hyperhidrosis using various methods. *Postepy Dermatol Alergol* 2009; **26**:501–5.
- Lowe NJ, Glaser DA, Eadie N *et al.* Botulinum toxin type A in the treatment of primary axillary hyperhidrosis: a 52-week multicenter double-blind, randomized, placebo-controlled study of efficacy and safety. *J Am Acad Dermatol* 2007; **56**:604–11.
- Naumann M, Lowe NJ. Botulinum toxin type A in treatment of bilateral primary axillary hyperhidrosis: randomised, parallel group, double blind, placebo controlled trial. *BMJ* 2001; **323**:596–9.
- Ohshima Y, Tamada Y, Yokozeki H *et al.* [The efficacy and safety of botulinum toxin type A in patients with primary axillary hyperhidrosis]. *Nishinihon J Dermatol* 2013; **75**:357–64 (in Japanese).
- Odderson IR. Long-term quantitative benefits of botulinum toxin type A in the treatment of axillary hyperhidrosis. *Dermatol Surg* 2002; **28**:480–3.
- Baumann L, Slezinger A, Halem M *et al.* Pilot study of the safety and efficacy of Myobloc (botulinum toxin type B) for treatment of axillary hyperhidrosis. *Int J Dermatol* 2005; **44**:418–24.
- Heckmann M, Ceballos-Baumann AO, Plewig G *et al.* Botulinum toxin A for axillary hyperhidrosis (excessive sweating). *N Engl J Med* 2001; **344**:488–93.
- Balzani A, Lupi F, Pezza M *et al.* Botulinum toxin type A in the treatment of axillary hyperhidrosis. *Dermatol Clin* 2001; **21**:117–20.
- Schnider P, Binder M, Kittler H *et al.* A randomized, double-blind, placebo-controlled trial of botulinum A toxin for severe axillary hyperhidrosis. *Br J Dermatol* 1999; **140**:677–80.
- Naumann M, Lowe NJ, Kumar CR *et al.* Botulinum toxin type A is a safe and effective treatment for axillary hyperhidrosis over 16 months: a prospective study. *Arch Dermatol* 2003; **139**:731–6.
- Heckmann M, Breit S, Ceballos-Baumann A *et al.* Side-controlled intradermal injection of botulinum toxin A in recalcitrant axillary hyperhidrosis. *J Am Acad Dermatol* 1999; **41**:987–90.
- Naver H, Swartling C, Aquilonius SM. Palmar and axillary hyperhidrosis treated with botulinum toxin: one-year clinical follow-up. *Eur J Neurol* 2000; **7**:55–62.
- Wakugawa M, Fujiyama M, Yoshida T *et al.* Half-side comparison study on the efficacy of botulinum toxin type A in patients with axillary hyperhidrosis. Presented at the 10th Congress of the European Academy of Dermatology and Venereology, Munich, Germany, 10–14 October 2001; 935–8.
- Ibrahim O, Kakar R, Bolotin D *et al.* The comparative effectiveness of suction-curettage and onabotulinumtoxin-A injections for the treatment of primary focal axillary hyperhidrosis: a randomized control trial. *J Am Acad Dermatol* 2013; **69**:88–95.
- Ottomann C, Blazek J, Hartmann B, Muehlberger T. [Liposuction curettage versus Botox for axillary hyperhidrosis. A prospective study of the quality of life]. *Chirurg* 2007; **78**:356–61 (in German).

- 34 Rompel R, Scholz S. Subcutaneous curettage vs. injection of botulinum toxin A for treatment of axillary hyperhidrosis. *J Eur Acad Dermatol Venerol* 2001; **15**:207–11.
- 35 Vakili D, Baker DM. Surgical micro-retrodermal axillary curettage (mRAC) effectively reduces axillary hyperhidrosis. *Br J Surg* 2016; **103**:12–13.
- 36 Baumann L, Slezinger A, Halem M *et al.* Double-blind, randomized, placebo-controlled pilot study of the safety and efficacy of Myobloc (botulinum toxin type B) for the treatment of palmar hyperhidrosis. *Dermatol Surg* 2005; **31**:263–70.
- 37 Lowe NJ, Yamauchi PS, Lask GP *et al.* Efficacy and safety of botulinum toxin type A in the treatment of palmar hyperhidrosis: a double-blind, randomized, placebo-controlled study. *Dermatol Surg* 2002; **28**:822–7.
- 38 Schnider P, Binder M, Auff E *et al.* Double-blind trial of botulinum A toxin for the treatment of focal hyperhidrosis of the palms. *Br J Dermatol* 1997; **136**:548–52.
- 39 Yamashita N, Shimizu H, Kawada M *et al.* Local injection of botulinum toxin A for palmar hyperhidrosis: usefulness and efficacy in relation to severity. *J Dermatol* 2008; **35**:325–9.
- 40 Glogau RG. Topically applied botulinum toxin type A for the treatment of primary axillary hyperhidrosis: results of a randomized, blinded, vehicle-controlled study. *Dermatol Surg* 2007; **33**:S76–80.
- 41 Mehrotra S, Schmith VD, Dumitrescu TP *et al.* Pharmacometrics-guided drug development of antihyperhidrosis agents. *J Clin Pharmacol* 2015; **55**:1256–67.
- 42 Hyun MY, Son IP, Lee Y *et al.* Efficacy and safety of topical glycopyrrolate in patients with facial hyperhidrosis: a randomized, multicentre, double-blinded, placebo-controlled, split-face study. *J Eur Acad Dermatol Venerol* 2015; **29**:278–82.
- 43 Baker D. Topical glycopyrrolate spray 2% reduces axillary hyperhidrosis to a similar extent as Botox injections. *Br J Dermatol* 2013; **169**:4.
- 44 Wolosker N, de Campos JR, Kauffman P *et al.* A randomized placebo-controlled trial of oxybutynin for the initial treatment of palmar and axillary hyperhidrosis. *J Vasc Surg* 2012; **55**:1696–700.
- 45 Costa Ada S Jr, Leão LE, Succi JE *et al.* Randomized trial – oxybutynin for treatment of persistent plantar hyperhidrosis in women after sympathectomy. *Clinics* 2014; **69**:101–5.
- 46 Schollhammer M, Brenaut E, Menard-Andivot N *et al.* Oxybutynin as a treatment for generalized hyperhidrosis: a randomized, placebo-controlled trial. *Br J Dermatol* 2015; **173**:1163–8.
- 47 Hund M, Sinkgraven R, Rzany B. [Randomized, placebo-controlled, double blind clinical trial for the evaluation of the efficacy and safety of oral methanthelinium bromide (Vagantin) in the treatment of focal hyperhidrosis]. *J Dtsch Dermatol Ges* 2004; **2**:343–9 (in German).
- 48 Müller C, Berensmeier A, Hamm H *et al.* Efficacy and safety of methantheline bromide (Vagantin) in axillary and palmar hyperhidrosis: results from a multicenter, randomized, placebo-controlled trial. *J Eur Acad Dermatol Venerol* 2013; **27**:1278–84.
- 49 Bechara FG, Sand M, Hoffmann K *et al.* Histological and clinical findings in different surgical strategies for focal axillary hyperhidrosis. *Dermatol Surg* 2008; **34**:1001–9.
- 50 Tronstad C, Helsing P, Tonseth KA *et al.* Tumescence suction curettage vs. curettage only for treatment of axillary hyperhidrosis evaluated by subjective and new objective methods. *Acta Derm Venerol* 2014; **94**:215–20.
- 51 Bechara FG, Sand M, Hoffmann K *et al.* Aggressive shaving after combined liposuction and curettage for axillary hyperhidrosis leads to more complications without further benefit. *Dermatol Surg* 2008; **34**:952–3.
- 52 Jemec B. Abrasio axillae in hyperhidrosis. *Scand J Plast Reconstr Surg* 1975; **9**:44–6.
- 53 Leclère FM, Moreno-Moraga J, Alcolea JM *et al.* Efficacy and safety of laser therapy on axillary hyperhidrosis after one year follow-up: a randomized blinded controlled trial. *Lasers Surg Med* 2015; **47**:173–9.
- 54 Bechara FG, Georgas D, Sand M *et al.* Effects of a long-pulsed 800-nm diode laser on axillary hyperhidrosis: a randomized controlled half-side comparison study. *Dermatol Surg* 2012; **38**:736–40.
- 55 Letada PR, Landers JT, Uebelhoefer NS *et al.* Treatment of focal axillary hyperhidrosis using a long-pulsed Nd:YAG 1064 nm laser at hair reduction settings. *J Drugs Dermatol* 2012; **11**:59–63.
- 56 Fatemi Naeini F, Abtahi-Naeini B, Pourazizi M *et al.* Fractionated microneedle radiofrequency for treatment of primary axillary hyperhidrosis: a sham control study. *Australas J Dermatol* 2015; **56**:279–84.
- 57 Glaser DA, Coleman WP 3rd, Fan LK *et al.* A randomized, blinded clinical evaluation of a novel microwave device for treating axillary hyperhidrosis: the dermatologic reduction in underarm perspiration study. *Dermatol Surg* 2012; **38**:185–91.
- 58 Nestor MS, Park H. Safety and efficacy of micro-focused ultrasound plus visualization for the treatment of axillary hyperhidrosis. *J Clin Aesthet Dermatol* 2014; **7**:14–21.

## Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

**Appendix S1** MEDLINE search strategy.

**Appendix S2** Risk-of-bias assessment results.

**Appendix S3** Studies that met the inclusion criteria for the review – database searches January 2016 (n = 155).

**Appendix S4** Flow diagram of the study selection process for conference proceedings and trial register searches.

**Appendix S5** Study details and results tables.

**Table S1** Basic study characteristics.

**Powerpoint S1** Journal Club Slide Set.



## Hyperhidrosis quality of life measures: review and patient perspective

Ros Wade, Julie Jones-Diette, Kath Wright, Alison M Layton & Nerys Woolacott

To cite this article: Ros Wade, Julie Jones-Diette, Kath Wright, Alison M Layton & Nerys Woolacott (2019) Hyperhidrosis quality of life measures: review and patient perspective, Journal of Dermatological Treatment, 30:3, 303-308, DOI: [10.1080/09546634.2018.1506080](https://doi.org/10.1080/09546634.2018.1506080)

To link to this article: <https://doi.org/10.1080/09546634.2018.1506080>

 View supplementary material 

 Published online: 12 Sep 2018.

 Submit your article to this journal 

 Article views: 298

 View related articles 

 View Crossmark data 

 Citing articles: 16 View citing articles 

ARTICLE



## Hyperhidrosis quality of life measures: review and patient perspective

Ros Wade<sup>a</sup> , Julie Jones-Diette<sup>a</sup>, Kath Wright<sup>a</sup>, Alison M. Layton<sup>b</sup> and Nerys Woolcott<sup>a</sup>

<sup>a</sup>Centre for Reviews and Dissemination, University of York, York, United Kingdom; <sup>b</sup>Harrogate and District NHS Foundation Trust, Harrogate, United Kingdom

### ABSTRACT

**Purpose:** To identify the tools that have been used to measure quality of life in hyperhidrosis research and obtain patient insight on commonly used tools.

**Methods:** Twelve databases were searched to identify studies that reported measuring quality of life or described a quality of life tool in the context of hyperhidrosis. Data on the use of the tools were tabulated and hyperhidrosis-specific and dermatology-specific measures were summarized. A workshop was held to obtain the patients' perspective on the most commonly used tools and the newly developed HidroQoL tool.

**Results:** One hundred and eighty-two studies were included in the review. Twenty-two quality of life tools were identified; two or more tools were often used in combination. The most commonly used tools were the Hyperhidrosis Disease Severity Scale, the Dermatology Quality of Life Index and the Hyperhidrosis Quality-of-Life Questionnaire. Patient advisors preferred the new HidroQoL tool, which was considered to be easy to complete and most relevant to hyperhidrosis patients.

**Conclusions:** There are several tools available for assessing quality of life in hyperhidrosis patients; disease specific measures are widely used and appear suitable. It is unclear which tool is the most reliable, although the HidroQoL tool was preferred by a small group of patient advisors.

### ARTICLE HISTORY

Received 14 May 2018  
Accepted 13 June 2018

### KEYWORDS

Hyperhidrosis; quality of life; dermatology; review

### Introduction

Hyperhidrosis is characterized by uncontrollable excessive and unpredictable sweating. Primary hyperhidrosis has no discernible cause and is thought to affect ~1% of the UK population (1). It most commonly involves the axillae, hands and feet, but may also involve other areas of the body. The symptoms of hyperhidrosis can significantly affect a patient's quality of life, and can lead to social embarrassment, loneliness, anxiety, and depression. Primary hyperhidrosis usually develops in childhood and adolescence. Teenagers may struggle to do schoolwork and exams, due to problems holding a pen and sweating ruining paperwork in classwork or in exams. Adults may find the condition affects employability. It may prevent individuals having personal relationships. The unpredictable and uncontrollable nature of the condition can make it very distressing for sufferers.

It is important that treatments used to manage hyperhidrosis symptoms not only reduce sweating, but also have a beneficial effect on patients' quality of life. Therefore, health related quality of life should be assessed in clinical studies.

### Objectives

To identify the tools that have been used to measure quality of life in hyperhidrosis research and obtain patient insight and perspective on commonly used tools.

### Materials and methods

This review was undertaken as part of a broader project assessing the clinical and cost-effectiveness of interventions for the

management of primary hyperhidrosis in secondary care (2). The protocol for the broader project was registered on PROSPERO (number CRD42015027803).

To identify all the tools used for the assessment of quality of life in hyperhidrosis research we adopted two literature search strategies. Twelve electronic databases (including MEDLINE, EMBASE, and PsycINFO) were searched in January 2016. The first search strategy was conducted as part of the broader project and combined relevant search terms for 'hyperhidrosis' with search terms for treatment types, for example, 'iontophoresis'. The second search strategy combined search terms for 'hyperhidrosis' with a recognized search filter for 'quality of life'. No date or language limits were applied. The 'quality of life' specific search strategy developed for Ovid MEDLINE is presented as [Supplementary Appendix S1](#), together with the full list of electronic databases searched. Clinical advisors were consulted for additional potentially relevant studies and the reference lists of relevant reviews were manually searched. An update search of MEDLINE was performed in March 2018 to check for new studies of the HidroQoL tool; none were found.

Two reviewers (RW and JJ-D) single screened titles and abstracts obtained through the search, with a sample of 10% of records double screened to confirm agreement between the reviewers; the level of agreement between reviewers was 96.2%. Full manuscripts of potentially relevant studies were obtained and independently screened by two reviewers (RW and JJ-D), using pre-defined eligibility criteria. Disagreements were resolved through discussion or consultation with a third reviewer.

All studies that reported measuring quality of life or described a quality of life tool in the context of primary hyperhidrosis were

**Table 1.** Frequency data for the use of quality of life measures in 182 studies of hyperhidrosis.

Quality of life measure (tool)	Acronym/abbreviation	Total number of studies in which reported
<b>Hyperhidrosis specific measures</b>		
Hyperhidrosis Disease Severity Scale (3,4)	HDSS	63
Hyperhidrosis Quality-of-Life Questionnaire (5)	HQLQ	31
Keller Hyperhidrosis Scale (7)	Keller, 2001	10
Hyperhidrosis Impact Questionnaire (8)	HHIQ	8
Disease-Specific Health-Related Questionnaire for Hyperhidrosis (6)	Amir (2000)	5
Hyperhidrosis Quality of Life Index (9–11)	HidroQoL	3
Hyperhidrosis Questionnaire (12)	HQ	2
<b>Dermatology specific measures</b>		
Dermatology Quality of Life Index (includes children's version CDLQI; 13)	DLQI	48
Skindex – Quality of life measure for people with skin disease (23)	Skindex	6
VQ-Dermato scale – A French language scoring instrument validated for chronic skin diseases (20)	VQ-Dermato scale	1
Freiburg Life Quality Assessment (21)	FLQA	1
Patient Benefit Index (22)	PBI	1
<b>Generic quality of life tools</b>		
Short Form 36 health status survey (24)	SF-36	13
Short Form 12 health status survey (25)	SF-12	7
Illness Intrusiveness Rating Scale (26)	IIRS	2
Leibowitz Social Anxiety Scale (27)	Liebowitz (1987)	1
Questionnaire of Quality of Life (adapted from the Caregiver Questionnaire; 28)	QQL	1
University of California Loneliness Scale (29)	UCLA V3	1
Nottingham Health Profile (30)	NHP	1
The Everyday Life Questionnaire (31)	EDLQ	1
State-Trait Anxiety Inventory (32)	STAI	1
International quality of life assessment (33)	EuroQoL 5D-5L (EQ-5D)	1

included. These studies were identified at the abstract screening stage or from the full papers ordered for the review of effectiveness. It is acknowledged that some papers excluded from the effectiveness review at the abstract stage may have mentioned quality of life in the full paper: such studies will have been missed. However, we consider that it is unlikely that any important quality of life tools have been missed, owing to the large number of studies screened.

Data extraction into Microsoft Excel comprised of details of the quality of life tool or tools used; whether the tool was disease specific for hyperhidrosis, disease specific for skin disease, or a generic quality of life tool; and any description of the validity of the tool was also extracted, where available. Data were extracted by a single reviewer and checked by a second reviewer (RW and JJ-D).

The included studies were not quality assessed as they were not necessarily evaluating the effectiveness of interventions, nor was the information extracted effectiveness data. While the COSMIN quality checklist suggests it could be useful when selecting a measurement instrument, it was found that it could not be readily used as it requires a high level of detailed information about how a tool was developed, far more than was available for this review; the studies found did not provide sufficient information to enable such a detailed assessment of methodological quality.

Data on the use of the quality of life tools were tabulated and hyperhidrosis-specific and dermatology-specific measures were summarized in a narrative synthesis. The aim of the review was to provide an overview of tools used in hyperhidrosis. As a formal validation of each tool was beyond the remit of this review no statistical analysis was undertaken.

## Results

The searches identified 337 publications in total, of which 182 studies were relevant for inclusion in the review. Twenty-two individual tools for measuring quality of life were identified, summarized in Table 1. Some studies reported using more than one tool, hence the total number of studies in which the tools were reported is 208.

A brief description of the hyperhidrosis and dermatology specific measures is presented below.

### Hyperhidrosis specific measures

#### Hyperhidrosis Disease Severity Scale (HDSS)

The HDSS was identified as the most commonly used tool; it was used in 63 studies in total, in both surgical and medical hyperhidrosis research. The HDSS was often used in combination with the Dermatology Life Quality Index (DLQI), with 18 studies using both tools.

The HDSS is a disease specific tool considered important for diagnostic use in clinical practice and for research to identify and quantify the severity of disease in patients with hyperhidrosis and also to assess treatment effects over time (3,4). The HDSS allows researchers to measure the impact hyperhidrosis has on those suffering from excessive sweating using a four-point scale:

1. My sweating is never noticeable and never interferes with my daily activities.
2. My sweating is tolerable but sometimes interferes with my daily activities.
3. My sweating is barely tolerable and frequently interferes with my daily activities.
4. My sweating is intolerable and always interferes with my daily activities.

The tool's simple design has raised questions of its value as a tool to measure patient reported quality of life and a consensus exercise by the Canadian Hyperhidrosis Advisory Committee selected the HDSS more as a measure of disease severity (3). However, an assessment of the validity and reliability of the HDSS found that HDSS score 4 weeks post treatment correlated well with the DLQI and relevant activity items from the Hyperhidrosis Impact Questionnaire (HHIQ;  $r = 0.35-0.77$ ;  $p = <.001$ ; 4).

#### Hyperhidrosis quality-of-Life Questionnaire (HQLQ)

The HQLQ was designed by De Campos and colleagues in 2003 as a disease specific tool to assess the effect of surgical interventions for patients with hyperhidrosis (5). The design built upon

the previous validation work of Amir and colleagues (6), described below, and tested the tool on a patient sample ( $n=378$ ) with the aim of replacing more generic quality of life measures.

The HQLQ questionnaire consists of a single question to start 'how would you rate your quality of life before and after treatment' and the patient is asked to enter a score between 1 (excellent/much better) and 5 (very poor/much worse). This is followed by twenty questions selected for relevance from the 35 items in the Amir questionnaire (6), again scored between 1 and 5. The final score for quality of life has a range from 20 (excellent/much better after surgery) to 100 (very poor/much worse after surgery). No validation or reliability statistics or cross-validation with other tools was reported.

#### **Keller Hyperhidrosis Scale**

The Keller Hyperhidrosis Scale was designed by Keller and colleagues in 2001 to measure preoperative and postoperative quality of life scores of patients receiving bilateral endoscopic thoracic sympathectomy for palmar and plantar hyperhidrosis (7). The tool measures quality of life on a scale of 0 (mild) to 10 (severe). The validation work compared patient scores against the Short Form 36 (SF-36) and validation work reported a strong level of reliability (Cronbach's  $\alpha=0.89$ ).

#### **Hyperhidrosis Impact Questionnaire (HHIQ)**

The HHIQ was designed by Teale and colleagues in 2002 to assess the impact of hyperhidrosis on the daily lives of patients and measure the effect of treatment (8). The development of the tool was industry funded and its relative popularity is predominantly an effect of its use in Allergan research trials. The design of the tool was informed by a review of the literature and interviews with key stakeholders (patients and physicians in the UK and Germany) and then a pilot study with the same stakeholders tested the validity and linguistic equivalence of the questionnaire (8). The questionnaire contained four sections (i) disease and treatment background, (ii) direct impact on medical and non-medical resource utilization, (iii) indirect impact on employment and productivity, and (iv) intangible impacts on emotional status. Forty one questions measured baseline impact of the disease with 10 further questions for follow up assessments. The final design of the HHIQ was validated against the Short Form 12 (SF-12) health survey and the DLQI using a population of 345 patients and 145 non-hyperhidrosis controls. A test-retest of the 10 follow up questions using a cohort of clinical patients found consistent reliability and responsiveness.

#### **Disease-Specific Health-Related Questionnaire for Hyperhidrosis**

This tool was designed and validated by Amir and colleagues in Israel for a patient population who were considering surgery for hyperhidrosis. The tool was designed to assist with clinical decision making and to measure the efficacy of surgical interventions on sweat reduction (6).

The Amir tool was designed with 35 questions separated into the five domains with a seven point Likert scale for each response, where a score of 6–7 indicated a very low quality of life, 3–5 a medium level of quality of life and 1–2 a high level of quality of life. The validation exercise found a high level of reliability (Cronbach's  $\alpha=0.84$ ). However, a limitation of the validation work noted by the authors is that only patients waiting for surgery were used in the survey and therefore may represent only patients whose symptoms were more severe (6). In addition, the tool was designed and validated in Israel and reported in studies conducted in Brazil, both countries have a very hot climate that

could have an impact on the patient population and subsequent patient reported outcome measures.

#### **Hyperhidrosis Quality of Life Index (HidroQoL)**

The HidroQoL is a recently developed tool, identified via publications describing its design and extensive validation (9–11).

The tool was developed as a disease specific aid to both clinical practice and research to assist with hyperhidrosis patient/clinician communication. In 2012, Kamudoni and colleagues recruited an online cohort of 71 patients from a number of social networking sites to participate in initial interviews (9). This led to the development of a pilot tool containing 47 questions answered using a six-point scale. Further work in 2015 (11) used modern test theory to examine differential item functioning. The second stage of validation involved a cross-sectional cohort of 595 patients who completed a number of questionnaires for comparison (HDSS, DLQI, and Skindex-17). The HidroQoL correlated well with the DLQI ( $r=0.6$ ,  $p<.01$ ) and HDSS ( $r=0.59$ ,  $p<.001$ ) and showed correlation to the Skindex-17 scale but to a lesser extent. Reliability, tested using baseline measures and a test-retest method, showed strong reproducibility (internal consistency, Cronbach's  $\alpha$  overall scale=0.89; test-retest reliability, intra-class correlation=0.93,  $p<.001$ ).

An online longitudinal study involved 260 patients completing the tool on three separate occasions; the results indicated that the tool was responsive at identifying slight changes or small responses to treatment over time.

#### **Hyperhidrosis Questionnaire (HQ)**

The design and validation of the HQ was described by Kuo and colleagues in 2004 (12). The tool's development was informed by a review of the literature, followed by interviews with patients, nursing staff and clinicians. The pilot questionnaire contained 34 questions answered using a scale of 1 (least disturbance) to 5 (most disturbance). The study included 85 patients suffering from a combination of plantar, palmar, axilla or generalized hyperhidrosis attending a thoracic surgery outpatient clinic in Southern Taiwan between April 2002 and March 2003. Internal reliability and construct validity was reported (Cronbach's  $\alpha=0.95$ , range 0.71–0.94 across domains), but no cross-validation with other scales was reported. The final questionnaire contained 29 questions across five domains: functional; psychological; social; affective; and physical.

#### **Dermatology specific measures**

##### **Dermatology Life Quality Index (DLQI) includes children's version (CDLQI)**

The DLQI was the most commonly used dermatology specific tool, used in 48 studies. As mentioned previously, it was often used alongside the HDSS in hyperhidrosis research. The DLQI is a concise tool (10 questions) often used in the management of chronic skin disorders (13). Developed and validated by Finlay and Khan (1994) to provide a patient centered method for comparison between different types of skin disease, the questionnaire records the impact the disease has on a patient's quality of life and the relative effectiveness of treatment (13).

A review of the DLQI in 2004 reported that repeatability, internal consistency and sensitivity to change have all been demonstrated for this tool and it has been cross-validated against a number of other dermatology tools, mainly for psoriasis and acne (14). However, more recently, detailed Rasch analysis has highlighted several problems with the scale, particularly when

combining DLQI scores for individuals with different types of skin condition (15).

#### **Skindex – Quality of life measure for people with skin disease**

The Skindex suite of tools includes the original Skindex questionnaire, Skindex-29, Skindex-17, and Skindex-16. The tool's development was based on findings from a review of the literature and focus group interviews with patients and clinicians to construct the initial framework for the ways in which patients are affected by skin disease. The original tool was a 61-question survey developed and validated by Chren and colleagues in 1996 (16), this was refined to a 29-item questionnaire (Skindex-29) to reduce completion time and improve the tool's evaluative properties (17). Further refinement resulted in a 16-item questionnaire (Skindex-16; 18) for use in longitudinal research to measure changes over time in patient quality of life in addition to reducing the tool to one page. The final version of the tool (Skindex-17) was created in 2006, using a response theory model to address issues such as response order and differential item functioning (19).

In a study of 201 patients, Skindex tool scores were reproducible after 72 h and were internally consistent (Cronbach's  $\alpha = 0.76-0.86$ ). Construct validity was also demonstrated. However, physicians' judgement of disease severity did not consistently correspond with Skindex scores and Skindex does not appear to have been cross-validated against other quality of life measures.

#### **VQ-Dermato scale – A French language scoring instrument validated for chronic skin diseases**

The VQ-Dermato scale was designed and validated by Grob and colleagues in 1999 to provide a French language dermatology-specific instrument for routine use to assess the quality of life of patients with 'chronic skin disorders' (20). The VQ-Dermato scale is a 28-item instrument developed from interviews with patients. The tool was validated on a population of 231 hospital and private practice patients in France suffering from chronic skin conditions. A strong correlation was reported between the VQ-Dermato scale and the SF-36 (Cronbach's  $\alpha = 0.67-0.88$ ).

#### **Freiburg Life Quality Assessment (FLQA)**

The FLQA was designed and validated as a set of dermatology-specific modules, the first module addressed the core issues of all skin diseases. The additional questions were more specific to distinct diseases. The tool was found to have satisfactory discriminatory power and validation data was published in 2004 (21).

#### **Patient Benefit Index (PBI)**

The PBI, developed by Augustin and colleagues in 2009, is an instrument used to identify patient reported needs and benefits of dermatology research and treatment. Assessment is a two-step process; the first to capture data on the patients' needs prior to treatment, followed by an assessment of improvement after treatment. The result is an index of patient benefit in response to treatment. The measure was validated in 2009 using a large cohort of patients ( $n = 500$ ) with many different skin diseases, including hyperhidrosis ( $n = 50$ ; 22).

#### **Patients' perspective**

The patients' perspective was collected to complement the narrative review of quality of life measures used in hyperhidrosis research. A workshop was held at Harrogate District Hospital with four patient advisors and one dermatologist (AML). All four

patients had moderate to severe hyperhidrosis for over 5 years; two patients had hyperhidrosis of the axilla and two had hyperhidrosis of the hand and axilla. Three patients were female and one was male and patients' ages ranged from their 20s to 50s. Prior to the workshop the patient advisors were sent copies of four quality of life tools: the three most commonly used tools (HDSS, DLQI, and HQLQ) and the newly developed HidroQoL tool, and asked to consider a short list of questions about the tools (see [Supplementary Appendix S2](#)). At the workshop the review of quality of life tools used in hyperhidrosis research was described and patients were asked to comment on the four tools.

All patient advisors agreed that the HidroQoL tool was superior to the other three tools. They commented that it covers everything important to patients with hyperhidrosis and is easy to complete. The DLQI was considered to be too general and too focused on the skin, with questions that were not applicable to hyperhidrosis patients. The HDSS was considered to be too basic and, depending on different situations, patients could easily fluctuate between an HDSS score of 2 or 3. More generally the patient advisors considered that measuring the actual amount of sweat produced (e.g. by gravimetry) was less important than measuring quality of life, and it should only be considered as a secondary outcome. They also stated that single measurements in time could give the wrong impression of the severity of hyperhidrosis and do not necessarily reflect the patient's overall condition. The patient advisors considered that the HidroQoL tool should be the primary outcome in future studies of interventions for hyperhidrosis.

## **Discussion**

The aim of this review was to identify the tools used to measure quality of life in hyperhidrosis research. The review identified a number of tools; the HDSS, the DLQI and the HQLQ were used more often than any other tool for measuring quality of life in hyperhidrosis research. The HDSS appears to have value for researchers assessing the clinical effectiveness of treatments for hyperhidrosis; it is often used to measure response to treatment. It is unclear from the literature what measures were used to design or validate the tool and it is not highly regarded as a comprehensive tool for measuring quality of life. The DLQI has a patient centered approach but it is criticized in the context of quality of life measures for hyperhidrosis for being too general and its inability to capture hyperhidrosis specific problems or concerns (11). UK and American studies commonly used the HDSS and DLQI in combination for both surgical and medical hyperhidrosis intervention studies. The HQLQ was designed specifically for surgical interventions for hyperhidrosis making it a popular choice for surgical studies although the majority of users were in Brazil where the tool was originally developed, with none of the studies being UK based.

Of interest is the new HidroQoL tool, developed by UK researchers as a scoring system with more focus on patient relevant measures than most quality of life tools used in hyperhidrosis research, for both research and clinical practice (10,11). This tool was not found in any studies assessing interventions for hyperhidrosis identified for the review although this may be because the tool is still relatively new.

In summary, there are a number of tools available for assessing quality of life in patients with hyperhidrosis. Disease specific measures are widely used and appear appropriate, although with the lack of standardization in method of development and validation it is not clear from this review which tool is the most

reliable. Some of the commonly used tools, such as the HDSS, appear to lack any form of published validation during development. The combined use of two or more tools is common, but again there is a lack of clear standardization for which combinations should be used or work best together. The type of intervention (surgical or medical) and geographical location may also be a factor in tool selection and it was not uncommon to find colleagues using the same tool. The HidroQoL is the most recent tool to be designed and validated for measuring the quality of life of patients with hyperhidrosis and was preferred by our small group of patient advisors.

## Conclusions

Health related quality of life should be a key outcome in future studies of interventions for hyperhidrosis. There are several tools available; disease specific measures are widely used and appear suitable. It is unclear which tool or tools are the most reliable for measuring quality of life in hyperhidrosis patients. The newly developed HidroQoL tool has been extensively validated and was preferred by a small group of patient advisors to this project. The HidroQoL tool should be tested alongside established tools, such as the HDSS and DLQI, to establish its reliability and patient/clinician acceptability in clinical practice and hyperhidrosis research.

## Acknowledgments

We would like to thank Dr Nick J Levell, Consultant Dermatologist at the Norfolk and Norwich University Hospital NHS Foundation Trust, Mr Gerard Stansby, Consultant Vascular Surgeon at the Newcastle upon Tyne Hospitals NHS Foundation Trust and Ms Julie Halford, Specialist Nurse at the Hampshire Clinic for clinical advice. We would also like to thank the four patient advisors for their comments and advice.

## Disclosure statement

The authors report no conflict of interest.

The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the Department of Health.

## Funding

This project was funded by the National Institute for Health Research Health Technology Assessment Programme (project number 14/211/02).

## ORCID

Ros Wade  <http://orcid.org/0000-0002-8666-8110>

## References

- Hyperhidrosis Support Group. What is hyperhidrosis? [Internet] Retrieved 21st July, 2016. Available from <http://www.hyperhidrosisuk.org/>
- Wade R, Rice S, Llewellyn A. Interventions for hyperhidrosis in secondary care: a systematic review and value-of-information analysis. *Health Technol Assess*. 2017;21:1–280.
- Solish N, Bertucci V, Dansereau A, et al. A comprehensive approach to the recognition, diagnosis, and severity-based treatment of focal hyperhidrosis: recommendations of the Canadian Hyperhidrosis Advisory Committee. *Dermatol Surg*. 2007;33:908–923.
- Kowalski JW, Eadie N, Dagget S, et al. Validity and reliability of the hyperhidrosis disease severity scale (HDSS). *J Am Acad Dermatol*. 2004;50:P51.
- de Campos JR, Kauffman P, Werebe Ede C, et al. Quality of life, before and after thoracic sympathectomy: report on 378 operated patients. *Ann Thorac Surg*. 2003;76:886–891.
- Amir M, Arish A, Weinstein Y, et al. Impairment in quality of life among patients seeking surgery for hyperhidrosis (excessive sweating): preliminary results. *Isr J Psych Relat Sci*. 2000;37:25–31.
- Keller S, Sekons D, Scher H, et al. A novel scale for assessing quality of life following bilateral endoscopic thoracic sympathectomy for palmar and plantar hyperhidrosis. Paper presented at the 4th International Symposium on Sympathetic Surgery, Tampere, Finland. 2001.
- Teale CW, Roberts G, Hamm H, et al. Development, validity, and reliability of the Hyperhidrosis Impact Questionnaire (HHIQ). *Qual Life Res*. 2002;11:702.
- Kamudoni P, Salek SS, Mueller B, et al. Hyperhidrosis quality of life index (Hidroqol<sup>®</sup>): a novel patient reported outcome measure in hyperhidrosis. *J Investigat Dermatol*. 2012;132:573.
- Salek S, Kamudoni P, Muller B. Interpretation of the constructs of Hyperhidrosis Quality of Life Index (HidroQoL) by two different cultures same language using classical and modern techniques. *Int J Clin Pharm*. 2015;37:266.
- Kamudoni P, Mueller B, Salek MS. The development and validation of a disease-specific quality of life measure in hyperhidrosis: the Hyperhidrosis Quality of Life Index (HidroQOL). *Qual Life Res*. 2015;24:1017–1027.
- Kuo CH, Yen M, Lin PC. Developing an instrument to measure quality of life of patients with hyperhidrosis. *J Nurs Res*. 2004;12:21–30.
- Finlay AY, Khan G. Dermatology Life Quality Index (DLQI)-a simple practical measure for routine clinical use. *Clin Exp Dermatol*. 1994;19:210–216.
- Lewis V, Finlay AY. 10 years experience of the Dermatology Life Quality Index (DLQI). *J Investigat Dermatol Symp Proc*. 2004;9:169–180.
- Twiss J, Meads DM, Preston EP, et al. Can we rely on the Dermatology Life Quality Index as a measure of the impact of psoriasis or atopic dermatitis? *J Invest Dermatol*. 2012;132:76–84.
- Chren M-M, Lasek RJ, Quinn LM, et al. Skindex, a quality-of-life measure for patients with skin disease: reliability, validity, and responsiveness. *J Invest Dermatol*. 1996;107:707–713.
- Chren M-M, Lasek RJ, Flocke SA, et al. Improved discriminative and evaluative capability of a refined version of Skindex, a quality-of-life instrument for patients with skin diseases. *Arch Dermatol*. 1997;133:1433–1440.
- Chren M-M, Lasek RJ, Sahay AP, et al. Measurement properties of Skindex-16: a brief quality-of-life measure for patients with skin diseases. *J Cutan Med Surg*. 2001;5:105–110.
- Nijsten TE, Sampogna F, Chren M-M, et al. Testing and reducing skindex-29 using Rasch analysis: Skindex-17. *J Investigat Dermatol*. 2006;126: 1244–1250.
- Grob JJ, Auquier P, Martin S, et al. Development and validation of a quality of life measurement for chronic skin

- disorders in French: VQ-Dermato. The RéseauEpidémiologie en Dermatologie. *Dermatology (Basel)*. 1999;199:213–222.
21. Augustin M, Lange S, Wenninger K, et al. Validation of a comprehensive Freiburg Life Quality Assessment (FLQA) core questionnaire and development of a threshold system. *Eur J Dermatol*. 2004;14:107–113.
  22. Augustin M, Radtke MA, Zschocke I, et al. The patient benefit index: a novel approach in patient-defined outcomes measurement for skin diseases. *Arch Dermatol Res*. 2009;301:561–571.
  23. Chren M-M. The Skindex instruments to measure the effects of skin disease on quality of life. *Dermatol Clin*. 2012;30:231–236.
  24. Stewart AL. *Measuring functioning and well-being: the medical outcomes study approach*. Durham: Duke University Press. 1992.
  25. Ware J, Kosinski M, Keller S. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med Care*. 1996;34:220–233.
  26. Devins GM. Illness intrusiveness and the psychosocial impact of lifestyle disruptions in chronic life-threatening disease. *Adv Ren Replace Ther*. 1994;1:251–263.
  27. Liebowitz MR. *Social phobia*. Basel: Karger Publishers. 1987.
  28. Schneider JW, Gurucharri LM, Gutierrez AL, et al. Health-related quality of life and functional outcome measures for children with cerebral palsy. *Dev Med Child Neurol*. 2001;43:601–608.
  29. Russell D, Peplau LA, Ferguson ML. Developing a measure of loneliness. *J Pers Assess*. 1978;42:290–294.
  30. Hunt SM, McEwen J, McKenna SP. Measuring health status: a new tool for clinicians and epidemiologists. *J R Coll Gen Pract*. 1985;35:185–188.
  31. Bullinger M, Kirchberger I, Von Steinbüchel N. Der Fragebogen Alltagsleben—ein Verfahren zur Erfassung der gesundheitsbezogenen Lebensqualität (The questionnaire everyday life – a procedure for recording the health-related quality of life). *Z Med Psychol*. 1993;3:121–131.
  32. Vazquez LD, Staples NL, Sears SF, et al. Psychosocial functioning of patients after endoscopic thoracic sympathectomy. *Eur J Cardiothorac Surg*. 2011;39:1018–1021.
  33. The EuroQoL G. EuroQoL—a new facility for the measurement of health-related quality of life. *Health Policy*. 1990;16:199–208.



## Health Technology Assessment

Volume 27 • Issue 29 • December 2023

ISSN 1366-5278

# Ablative and non-surgical therapies for early and very early hepatocellular carcinoma: a systematic review and network meta-analysis

*Ros Wade, Emily South, Sumayya Anwer, Sahar Sharif-Hurst, Melissa Harden,  
Helen Fulbright, Robert Hodgson, Sofia Dias, Mark Simmonds, Ian Rowe, Patricia Thornton  
and Alison Eastwood*





# Ablative and non-surgical therapies for early and very early hepatocellular carcinoma: a systematic review and network meta-analysis

Ros Wade<sup>1\*</sup>, Emily South<sup>1</sup>, Sumayya Anwer<sup>1</sup>,  
Sahar Sharif-Hurst<sup>1</sup>, Melissa Harden<sup>2</sup>, Helen Fulbright<sup>2</sup>,  
Robert Hodgson<sup>3</sup>, Sofia Dias<sup>4</sup>, Mark Simmonds<sup>3</sup>,  
Ian Rowe<sup>5</sup>, Patricia Thornton<sup>6</sup> and Alison Eastwood<sup>7</sup>

<sup>1</sup>Research Fellow, Centre for Reviews and Dissemination, University of York, Heslington, UK

<sup>2</sup>Information Specialist, Centre for Reviews and Dissemination, University of York, Heslington, UK

<sup>3</sup>Senior Research Fellow, Centre for Reviews and Dissemination, University of York, Heslington, UK

<sup>4</sup>Professor in Health Technology Assessment, Centre for Reviews and Dissemination, University of York, Heslington, UK

<sup>5</sup>Honorary Consultant Hepatologist, Leeds Teaching Hospitals NHS Trust

<sup>6</sup>Patient Collaborator

<sup>7</sup>Professor of Research, Centre for Reviews and Dissemination, University of York, Heslington, UK

\*Corresponding author

## Disclosure of interests

**Full disclosure of interests:** Completed ICMJE forms for all authors, including all related interests, are available in the toolkit on the NIHR Journals Library report publication page at <https://doi.org/10.3310/GK5221>.

**Primary conflicts of interest:** None.

**Rider on responsibility for report:** The views expressed in this report are those of the authors and not necessarily those of the NIHR HTA Programme. Any errors are the responsibility of the authors.

Published December 2023

DOI: 10.3310/GK5221

This report should be referenced as follows:

Wade R, South E, Anwer S, Sharif-Hurst S, Harden M, Fulbright H, *et al*. Ablative and non-surgical therapies for early and very early hepatocellular carcinoma: a systematic review and network meta-analysis. *Health Technol Assess* 2023;27(29). <https://doi.org/10.3310/GK5221>



# Health Technology Assessment

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 3.6

Launched in 1997, *Health Technology Assessment* (HTA) has an impact factor of 3.6 and is ranked 32nd (out of 105 titles) in the 'Health Care Sciences & Services' category of the Clarivate 2022 Journal Citation Reports (Science Edition). It is also indexed by MEDLINE, CINAHL (EBSCO Information Services, Ipswich, MA, USA), Embase (Elsevier, Amsterdam, the Netherlands), NCBI Bookshelf, DOAJ, Europe PMC, the Cochrane Library (John Wiley & Sons, Inc., Hoboken, NJ, USA), INAHTA, the British Nursing Index (ProQuest LLC, Ann Arbor, MI, USA), Ulrichsweb (ProQuest LLC, Ann Arbor, MI, USA) and the Science Citation Index Expanded (Clarivate, Philadelphia, PA, USA).

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) ([www.publicationethics.org/](http://www.publicationethics.org/)).

Editorial contact: [journals.library@nhr.ac.uk](mailto:journals.library@nhr.ac.uk)

The full HTA archive is freely available to view online at [www.journalslibrary.nhr.ac.uk/hta](http://www.journalslibrary.nhr.ac.uk/hta).

## Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

## HTA programme

Health Technology Assessment (HTA) research is undertaken where some evidence already exists to show that a technology can be effective and this needs to be compared to the current standard intervention to see which works best. Research can evaluate any intervention used in the treatment, prevention or diagnosis of disease, provided the study outcomes lead to findings that have the potential to be of direct benefit to NHS patients. Technologies in this context mean any method used to promote health; prevent and treat disease; and improve rehabilitation or long-term care. They are not confined to new drugs and include any intervention used in the treatment, prevention or diagnosis of disease.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

## This report

The research reported in this issue of the journal was funded by the HTA programme as project number NIHR131224. The contractual start date was in January 2021. The draft report began editorial review in February 2022 and was accepted for publication in August 2022. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health and Care Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, the HTA programme or the Department of Health and Social Care. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, the HTA programme or the Department of Health and Social Care.

Copyright © 2023 Wade *et al.* This work was produced by Wade *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health and Social Care. This is an Open Access publication distributed under the terms of the Creative Commons Attribution CC BY 4.0 licence, which permits unrestricted use, distribution, reproduction and adaptation in any medium and for any purpose provided that it is properly attributed. See: <https://creativecommons.org/licenses/by/4.0/>. For attribution the title, original author(s), the publication source – NIHR Journals Library, and the DOI of the publication must be cited.

Published by the NIHR Journals Library ([www.journalslibrary.nhr.ac.uk](http://www.journalslibrary.nhr.ac.uk)), produced by Newgen Digitalworks Pvt Ltd, Chennai, India ([www.newgen.co](http://www.newgen.co)).

## NIHR Journals Library Editor-in-Chief

---

**Dr Cat Chatfield** Director of Health Services Research UK

## NIHR Journals Library Editors

---

**Professor Andrée Le May** Chair of NIHR Journals Library Editorial Group (HSDR, PGfAR, PHR journals) and Editor-in-Chief of HSDR, PGfAR, PHR journals

**Dr Peter Davidson** Interim Chair of HTA and EME Editorial Board. Consultant Advisor, School of Healthcare Enterprise and Innovation, University of Southampton, UK

**Professor Matthias Beck** Professor of Management, Cork University Business School, Department of Management and Marketing, University College Cork, Ireland

**Dr Tessa Crilly** Director, Crystal Blue Consulting Ltd, UK

**Dr Eugenia Cronin** Consultant in Public Health, Delta Public Health Consulting Ltd, UK

**Ms Tara Lamont** Senior Adviser, School of Healthcare Enterprise and Innovation, University of Southampton, UK

**Dr Catriona McDaid** Reader in Trials, Department of Health Sciences, University of York, UK

**Professor William McGuire** Professor of Child Health, Hull York Medical School, University of York, UK

**Professor Geoffrey Meads** Emeritus Professor of Wellbeing Research, University of Winchester, UK

**Professor James Raftery** Professor of Health Technology Assessment, School of Healthcare Enterprise and Innovation, University of Southampton, UK

**Dr Rob Riemsma** Consultant Advisor, School of Healthcare Enterprise and Innovation, University of Southampton, UK

**Professor Helen Roberts** Professor of Child Health Research, Child and Adolescent Mental Health, Palliative Care and Paediatrics Unit, Population Policy and Practice Programme, UCL Great Ormond Street Institute of Child Health, London, UK

**Professor Jonathan Ross** Professor of Sexual Health and HIV, University Hospital Birmingham, UK

**Professor Helen Snooks** Professor of Health Services Research, Institute of Life Science, College of Medicine, Swansea University, UK

Please visit the website for a list of editors: [www.journalslibrary.nihr.ac.uk/about/editors](http://www.journalslibrary.nihr.ac.uk/about/editors)

**Editorial contact:** [journals.library@nihr.ac.uk](mailto:journals.library@nihr.ac.uk)

# Abstract

## Ablative and non-surgical therapies for early and very early hepatocellular carcinoma: a systematic review and network meta-analysis

Ros Wade<sup>1\*</sup>, Emily South<sup>1</sup>, Sumayya Anwer<sup>1</sup>, Sahar Sharif-Hurst<sup>1</sup>,  
Melissa Harden<sup>2</sup>, Helen Fulbright<sup>2</sup>, Robert Hodgson<sup>3</sup>, Sofia Dias<sup>4</sup>,  
Mark Simmonds<sup>3</sup>, Ian Rowe<sup>5</sup>, Patricia Thornton<sup>6</sup> and  
Alison Eastwood<sup>7</sup>

<sup>1</sup>Research Fellow, Centre for Reviews and Dissemination, University of York, Heslington, UK

<sup>2</sup>Information Specialist, Centre for Reviews and Dissemination, University of York, Heslington, UK

<sup>3</sup>Senior Research Fellow, Centre for Reviews and Dissemination, University of York, Heslington, UK

<sup>4</sup>Professor in Health Technology Assessment, Centre for Reviews and Dissemination, University of York, Heslington, UK

<sup>5</sup>Honorary Consultant Hepatologist, Leeds Teaching Hospitals NHS Trust

<sup>6</sup>Patient Collaborator

<sup>7</sup>Professor of Research, Centre for Reviews and Dissemination, University of York, Heslington, UK

\*Corresponding author [ros.wade@york.ac.uk](mailto:ros.wade@york.ac.uk)

**Background:** A wide range of ablative and non-surgical therapies are available for treating small hepatocellular carcinoma in patients with very early or early-stage disease and preserved liver function.

**Objective:** To review and compare the effectiveness of all current ablative and non-surgical therapies for patients with small hepatocellular carcinoma ( $\leq 3$  cm).

**Design:** Systematic review and network meta-analysis.

**Data sources:** Nine databases (March 2021), two trial registries (April 2021) and reference lists of relevant systematic reviews.

**Review methods:** Eligible studies were randomised controlled trials of ablative and non-surgical therapies, versus any comparator, for small hepatocellular carcinoma. Randomised controlled trials were quality assessed using the Cochrane Risk of Bias 2 tool and mapped. The comparative effectiveness of therapies was assessed using network meta-analysis. A threshold analysis was used to identify which comparisons were sensitive to potential changes in the evidence. Where comparisons based on randomised controlled trial evidence were not robust or no randomised controlled trials were identified, a targeted systematic review of non-randomised, prospective comparative studies provided additional data for repeat network meta-analysis and threshold analysis. The feasibility of undertaking economic modelling was explored. A workshop with patients and clinicians was held to discuss the findings and identify key priorities for future research.

**Results:** Thirty-seven randomised controlled trials (with over 3700 relevant patients) were included in the review. The majority were conducted in China or Japan and most had a high risk of bias or some risk of bias concerns. The results of the network meta-analysis were uncertain for most comparisons. There was evidence that percutaneous ethanol injection is inferior to radiofrequency ablation for overall survival (hazard ratio 1.45, 95% credible interval 1.16 to 1.82), progression-free survival (hazard ratio 1.36, 95% credible interval 1.11 to 1.67), overall recurrence (relative risk 1.19, 95% credible interval

1.02 to 1.39) and local recurrence (relative risk 1.80, 95% credible interval 1.19 to 2.71). Percutaneous acid injection was also inferior to radiofrequency ablation for progression-free survival (hazard ratio 1.63, 95% credible interval 1.05 to 2.51). Threshold analysis showed that further evidence could plausibly change the result for some comparisons. Fourteen eligible non-randomised studies were identified ( $n \geq 2316$ ); twelve had a high risk of bias so were not included in updated network meta-analyses. Additional non-randomised data, made available by a clinical advisor, were also included ( $n = 303$ ). There remained a high level of uncertainty in treatment rankings after the network meta-analyses were updated. However, the updated analyses suggested that microwave ablation and resection are superior to percutaneous ethanol injection and percutaneous acid injection for some outcomes. Further research on stereotactic ablative radiotherapy was recommended at the workshop, although it is only appropriate for certain patient subgroups, limiting opportunities for adequately powered trials.

**Limitations:** Many studies were small and of poor quality. No comparative studies were found for some therapies.

**Conclusions:** The existing evidence base has limitations; the uptake of specific ablative therapies in the United Kingdom appears to be based more on technological advancements and ease of use than strong evidence of clinical effectiveness. However, there is evidence that percutaneous ethanol injection and percutaneous acid injection are inferior to radiofrequency ablation, microwave ablation and resection.

**Study registration:** PROSPERO CRD42020221357.

**Funding:** This award was funded by the National Institute for Health and Care Research (NIHR) Health Technology Assessment (HTA) programme (NIHR award ref: NIHR131224) and is published in full in *Health Technology Assessment*; Vol. 27, No. 29. See the NIHR Funding and Awards website for further award information.

# Contents

List of tables	xi
List of figures	xiii
List of abbreviations	xv
List of supplementary material	xvii
Plain language summary	xix
Scientific summary	xxi
<b>Chapter 1</b> Background	<b>1</b>
<b>Chapter 2</b> Aim and objectives	<b>3</b>
<b>Chapter 3</b> Methods	<b>5</b>
Systematic review of randomised controlled trials	5
<i>Search strategy for identification of randomised controlled trials</i>	5
<i>Inclusion criteria</i>	6
<i>Study selection and data extraction</i>	7
<i>Critical appraisal</i>	8
Network meta-analysis	8
<i>Included data</i>	8
<i>Network meta-analysis</i>	8
Threshold analysis	9
Systematic review of non-randomised evidence	9
<i>Search strategy for identification of non-RCTs</i>	10
<i>Inclusion criteria</i>	11
<i>Study selection and data extraction</i>	12
<i>Critical appraisal</i>	12
Updated network meta-analysis	12
Updated threshold analysis	12
<b>Chapter 4</b> Results	<b>13</b>
Systematic review of RCTs	13
<i>Characteristics of RCTs included in the review</i>	13
<i>Quality of RCTs included in the review</i>	13
<i>Results of RCTs included in the review</i>	16
<i>Ongoing trials</i>	25
Network meta-analysis results	26
<i>Overall survival</i>	26
<i>Progression-free survival</i>	28
<i>Overall recurrence</i>	30
<i>Local recurrence</i>	32
Threshold analysis of RCT networks	35
<i>Overall survival</i>	35
<i>Progression-free survival</i>	36

## CONTENTS

Overall recurrence	38
Local recurrence	38
Systematic review of non-randomised evidence	38
Characteristics of non-randomised studies included in the review	41
Quality of non-randomised studies included in the review	43
Results of non-randomised studies included in the review	45
Ongoing trials	47
Updated network meta-analyses using RCT and non-RCT evidence	47
Overall survival	48
Progression-free survival	49
Local recurrence	52
Updated threshold analysis	53
Overall survival	53
Progression-free survival	55
Local recurrence	55
<b>Chapter 5 Feasibility of economic modelling</b>	<b>61</b>
Approach	61
Cost-effectiveness review	61
Review methods	61
Results	62
Model scope and availability of comparative data	66
Model structure and clinical data availability	67
Utilities and quality of life	68
Resource use and costs	68
Value of information	68
<b>Chapter 6 Patient and public involvement</b>	<b>69</b>
Aim	69
Methods	69
Results	69
Discussion and conclusions	70
Reflective/critical perspective	70
<b>Chapter 7 Workshop</b>	<b>71</b>
<b>Chapter 8 Discussion</b>	<b>75</b>
Summary of findings	75
Strengths and limitations	77
<b>Chapter 9 Conclusions</b>	<b>79</b>
Implications for practice	79
Recommendations for research	79
<b>Acknowledgements</b>	<b>81</b>
<b>References</b>	<b>83</b>
<b>Appendix 1 Search strategies</b>	<b>115</b>
<b>Appendix 2 Studies excluded at full paper stage with rationale (randomised controlled trial searches)</b>	<b>133</b>

<b>Appendix 3</b> Risk of bias assessment results (randomised controlled trials)	<b>137</b>
<b>Appendix 4</b> Characteristics and results of randomised controlled trials included in the review	<b>141</b>
<b>Appendix 5</b> Studies excluded at full paper stage with rationale (non-randomised controlled trial searches)	<b>161</b>
<b>Appendix 6</b> Characteristics and results of non-randomised studies included in the review	<b>167</b>



# List of tables

<b>TABLE 1</b> RCTs included in the systematic review	<b>15</b>
<b>TABLE 2</b> Number of RCTs making each comparison	<b>17</b>
<b>TABLE 3</b> Risk of bias assessment results (RCTs)	<b>18</b>
<b>TABLE 4</b> Radiofrequency ablation vs. PEI – OS	<b>20</b>
<b>TABLE 5</b> Table of potentially relevant ongoing RCTs	<b>25</b>
<b>TABLE 6</b> Mean and median treatment ranks for the FE model, with corresponding 95% Crls for OS, sorted by mean rank out of 11 treatments	<b>28</b>
<b>TABLE 7</b> Mean and median ranks for the FE model, with the corresponding 95% Crls for PFS, sorted by mean rank out of five treatments	<b>31</b>
<b>TABLE 8</b> Mean and median ranks for the FE model, with corresponding 95% Crls for overall recurrence, sorted by mean rank, out of seven treatments	<b>34</b>
<b>TABLE 9</b> Mean and median ranks for the FE model, with corresponding 95% Crls for local recurrence, sorted by mean rank, out of nine treatments	<b>36</b>
<b>TABLE 10</b> Non-RCTs included in the systematic review	<b>42</b>
<b>TABLE 11</b> Number of non-RCTs making each comparison	<b>43</b>
<b>TABLE 12</b> Risk of bias assessment results (non-RCTs)	<b>44</b>
<b>TABLE 13</b> Table of potentially relevant ongoing RCTs (identified from non-RCT searches)	<b>47</b>
<b>TABLE 14</b> Mean and median treatment ranks for the FE model, with corresponding 95% Crls for OS, sorted by mean rank out of 12 treatments	<b>50</b>
<b>TABLE 15</b> Mean and median ranks, with corresponding 95% Crls for PFS for the FE model, sorted by mean rank	<b>52</b>
<b>TABLE 16</b> Mean and median ranks, with corresponding 95% Crls, for local recurrence, sorted by mean rank, for the FE model	<b>55</b>
<b>TABLE 17</b> Data extraction: cost-effectiveness review	<b>63</b>



# List of figures

<b>FIGURE 1</b>	Flow diagram of the study selection process (RCTs)	14
<b>FIGURE 2</b>	Network diagram for OS	27
<b>FIGURE 3</b>	Plot of HRs for OS compared with RFA for the FE model	28
<b>FIGURE 4</b>	Rank plot for OS for the FE model	29
<b>FIGURE 5</b>	Network diagram for PFS	29
<b>FIGURE 6</b>	Plot of HRs for PFS compared with RFA for the FE model	30
<b>FIGURE 7</b>	Rank plot for PFS for the FE model	31
<b>FIGURE 8</b>	Network diagram for overall recurrence	32
<b>FIGURE 9</b>	Plot of RRs for overall recurrence compared with RFA for the FE model	33
<b>FIGURE 10</b>	Rank plot for overall recurrence for the FE model	33
<b>FIGURE 11</b>	Network diagram for local recurrence	34
<b>FIGURE 12</b>	Plot of RRs for local recurrence compared with RFA for the FE model	35
<b>FIGURE 13</b>	Rank plot for local recurrence for the FE model	36
<b>FIGURE 14</b>	Forest plot for threshold analysis results for OS	37
<b>FIGURE 15</b>	Forest plot for the threshold analysis for PFS	37
<b>FIGURE 16</b>	Forest plot for threshold analysis results for overall recurrence	39
<b>FIGURE 17</b>	Forest plot for results of threshold analysis for local recurrence	40
<b>FIGURE 18</b>	Flow diagram of the study selection process (non-RCTs)	41
<b>FIGURE 19</b>	Network diagram for OS	48
<b>FIGURE 20</b>	Plot of HRs for OS compared with RFA for the FE model	49
<b>FIGURE 21</b>	Rank plot for OS for the FE model	50
<b>FIGURE 22</b>	Network diagram for PFS	51
<b>FIGURE 23</b>	Plot of HRs for PFS compared with RFA for the FE model	51
<b>FIGURE 24</b>	Rank plot for PFS for the FE model	52

<b>FIGURE 25</b>	Network diagram for local recurrence	<b>53</b>
<b>FIGURE 26</b>	Plot of RRs for local recurrence compared with RFA for the FE model	<b>54</b>
<b>FIGURE 27</b>	Rank plot for local recurrence for the FE model	<b>54</b>
<b>FIGURE 28</b>	Forest plot for threshold analysis results for OS for the updated NMA	<b>56</b>
<b>FIGURE 29</b>	Forest plot for the threshold analysis for PFS for the updated NMA	<b>57</b>
<b>FIGURE 30</b>	Forest plot for results of threshold analysis for local recurrence for the updated NMA	<b>58</b>
<b>FIGURE 31</b>	Flow diagram of the study selection process (cost-effectiveness review)	<b>62</b>

## List of abbreviations

BCLC	Barcelona Clinic Liver Cancer	PEI	percutaneous ethanol injection
CI	confidence interval	PFS	progression-free survival
CRD	Centre for Reviews and Dissemination	PSM	partition survival model
CrI	credible interval	RCT	randomised controlled trial
CT	computerised tomography	RE	random-effects
DIC	deviance information criteria	RFA	radiofrequency ablation
ECT	electrochemotherapy	ROB	risk of bias
FE	fixed-effect	RR	relative risk
HBV	hepatitis B virus	SABR	stereotactic ablative radiotherapy (this is the same technology as SBRT, for simplicity SABR is used throughout this report)
HCC	hepatocellular carcinoma		
HCV	hepatitis C virus		
HIFU	high-intensity focused ultrasound	SBRT	stereotactic body radiotherapy (this is the same technology as SABR, see comment above)
HR	hazard ratio		
HRQOL	health-related quality of life		
ICU	intensive care unit	SD	standard deviation
IRE	irreversible electroporation	SE	standard error
KM	Kaplan–Meier	SIRT	selective internal radiation therapy
MWA	microwave ablation		
NICE	National Institute for Health and Care Excellence	TACE	transarterial chemoembolisation
NMA	network meta-analysis	TTP	time to progression
OS	overall survival	VOI	value of information
PAI	percutaneous acetic acid injection		



# List of supplementary material

**Report Supplementary Material 1** Full search strategies

**Report Supplementary Material 2** Schoenfeld residuals

**Report Supplementary Material 3** Reasons for inclusion and exclusion of RCTs in the NMA

**Report Supplementary Material 4** Data tables for NMAs

**Report Supplementary Material 5** NMA results

Supplementary material can be found on the NIHR Journals Library report page (<https://doi.org/10.3310/GK5221>).

Supplementary material has been provided by the authors to support the report and any files provided at submission will have been seen by peer reviewers, but not extensively reviewed. Any supplementary material provided at a later stage in the process may not have been peer reviewed.



## Plain language summary

**H**epatocellular carcinoma is the most common type of primary liver cancer. There are a range of different treatments available for patients with early hepatocellular carcinoma. We looked for clinical trials in patients with small tumours (up to 3 cm) that compared different treatments. We brought together and analysed the results of these trials to see which treatments were most effective in terms of survival, progression, side effects and quality of life.

Overall, the evidence has limitations; many trials had few patients and were of poor quality. Most were from China or Japan, where the common causes of liver disease and treatments available differ from those in the United Kingdom. The results of our analyses were very uncertain so we cannot be sure which treatment is the best overall.

We did find that three treatments – radiofrequency ablation, microwave ablation and surgery – were generally more effective than percutaneous ethanol injection and percutaneous acid injection. There was not enough evidence to be certain which treatment was better when radiofrequency ablation was compared with laser ablation, microwave ablation, proton beam therapy or surgery. We found only poor-quality, non-randomised trials on high-intensity focused ultrasound, cryoablation and irreversible electroporation. There was very little evidence on treatments that combined radiofrequency ablation with other therapies. We found no studies that compared electrochemotherapy, histotripsy, stereotactic ablative radiotherapy or wider radiotherapy techniques with other treatments. Only two studies reported data on quality of life or patient satisfaction.

We discussed the findings with patients and clinical experts. Stereotactic ablative radiotherapy was highlighted as a treatment that requires further research; however, it is only appropriate for certain subgroups of patients. Feasibility studies could inform future clinical trials by exploring issues such as whether patients are willing to take part in a trial or find the treatments acceptable.



# Scientific summary

## Background

Hepatocellular carcinoma (HCC) is the most common type of primary liver cancer. Around one-third of people with cirrhosis go on to develop HCC. The prognosis of symptomatic HCC is poor, so the National Institute for Health and Care Excellence recommends that patients with cirrhosis are monitored for early HCC with six-monthly ultrasound scans.

Patients with early HCC and good liver function can be offered surgical or non-surgical interventions with curative intent. However, liver resection is not always possible due to the location of the tumour, poor liver function or portal hypertension, and liver transplantation is limited by availability. Therefore, ablative or non-surgical therapies are frequently used for treating early HCC, including microwave ablation (MWA) and radiofrequency ablation (RFA). There has been no definitive assessment of these therapies.

## Objectives

The aim of this project was to evaluate and compare the effectiveness of ablative and non-surgical therapies for patients with small HCC.

The key objectives were to:

- systematically identify all randomised controlled trials (RCTs) of ablative and non-surgical therapies for HCC
- evaluate their quality and applicability to UK populations
- determine the comparative effectiveness of therapies using network meta-analysis (NMA)
- where the evidence base is insufficient, supplement the RCT evidence with high-quality, non-randomised, prospective comparative studies
- identify priority areas where additional high-quality evidence is required (in collaboration with patients and clinicians)
- assess whether future economic analysis would be feasible and worthwhile.

## Methods

### *Systematic review of randomised controlled trials*

Nine databases (including MEDLINE, Embase, CENTRAL, Science Citation Index) were searched for RCTs and systematic reviews published from 2000 to March 2021. Two trial registries were searched in April 2021 to identify ongoing and unpublished RCTs. The reference lists of relevant systematic reviews were checked and clinical advisors were consulted.

Randomised controlled trials of patients with HCC up to 3 cm in size (or data on a subgroup(s) of patients with tumours  $\leq$  3 cm) were eligible for inclusion. Any ablative or non-surgical therapy was eligible, including:

- RFA
- MWA
- laser ablation
- high-intensity focused ultrasound (HIFU)

- cryoablation
- percutaneous ethanol injection (PEI)
- percutaneous acetic acid injection (PAI)
- irreversible electroporation (IRE)
- transarterial chemoembolisation (TACE)
- transarterial embolisation
- selective internal radiation therapy
- electrochemotherapy (ECT)
- histotripsy
- stereotactic ablative radiotherapy [SABR; the term stereotactic body radiotherapy (SBRT) is also used for this technology]
- wider radiotherapy techniques.

Any comparator was eligible, except a different method of undertaking the same intervention. Outcomes of interest were overall survival (OS), progression-free survival (PFS), time to progression (TTP), serious adverse events (AEs), intervention-specific AEs and quality of life.

Titles and abstracts were screened by one reviewer, with 10% checked by another reviewer. Full texts were screened by two reviewers independently. Data extraction was checked by a second reviewer. Risk of bias (RoB) was assessed using the Cochrane RoB 2 tool. When studies did not report hazard ratios (HRs) and their variances, Kaplan–Meier data were extracted.

### ***Network meta-analysis***

After mapping the identified RCTs, NMAs were conducted for four outcomes: OS, PFS, overall recurrence and local recurrence. They were conducted in a Bayesian framework using Markov chain Monte Carlo techniques. The NMAs were used to assess and rank interventions by comparative effectiveness.

### ***Threshold analysis***

Threshold analysis was conducted at the contrast level to examine the impact of potential changes to the evidence on each treatment contrast. Results of the analysis were used to identify treatment comparisons which lacked robust RCT evidence and where non-randomised evidence should be sought for further review.

### ***Systematic review of non-randomised evidence***

A second systematic review of non-randomised evidence was undertaken. This review included studies of comparisons where additional evidence could plausibly change the NMA conclusions, as identified by the threshold analysis. Four databases were searched in August 2021 for studies that compared the selected interventions (RFA, MWA and laser ablation), either with each other or with resection.

The databases were also searched in July 2021 for interventions that the advisory group identified as being of particular interest and where there was no RCT evidence: HIFU, cryoablation, IRE, ECT, histotripsy, SABR and wider radiotherapy techniques.

Prospective non-randomised comparative trials of patients with HCC up to 3 cm (or data on a subgroup(s) of such patients) were eligible. The outcomes of interest were OS, PFS, TTP and quality of life.

Methods of screening and data extraction were the same as outlined above. A validity assessment tool for non-randomised trials was developed.

### **Updated network meta-analysis and threshold analysis**

Where the non-randomised trials were of sufficient quality, the NMAs were repeated after pooling (without any adjustments) the non-randomised evidence with the RCT evidence, to assess whether estimates were improved. A threshold analysis was conducted on the updated NMA results to explore robustness and sensitivity to bias of the new results.

## **Results**

### **Systematic review of randomised controlled trial results**

Thirty-seven RCTs were included. Most were small, with sample sizes ranging from 30 to 308 patients. The majority of RCTs were conducted in China or Japan. The most frequently assessed therapy was RFA. The majority of RCTs assessed OS, PFS/disease-free survival and/or recurrence, along with response and AEs. One RCT assessed patient satisfaction. The RoB judgement was low for 9 RCTs, high for 12 RCTs and some concerns for 14 RCTs (two RCTs that reported no relevant outcomes were not assessed).

For many comparisons, data were limited. Based on a narrative synthesis, RFA appears to be better than both PEI and PAI in terms of OS, PFS and recurrence, although AEs were more frequent after RFA. PAI appears to have similar effectiveness to PEI. For RFA versus resection, results were inconsistent, with some RCTs favouring RFA and some resection; AEs were more frequent after resection. Data from RCTs comparing RFA with MWA, laser ablation or proton beam therapy were limited. RCTs assessing RFA in combination with other treatments were also limited by small sample sizes. AEs were reported inconsistently. There was no RCT evidence for HIFU, cryoablation, IRE, ECT, histotripsy, SABR or wider radiotherapy techniques.

### **Network meta-analysis and threshold analysis results**

The treatment rankings from the NMAs were very uncertain for all four outcomes (OS, PFS, overall and local recurrence). There was no meaningful difference in effectiveness for many of the treatment comparisons.

There was evidence that PEI is worse than RFA for OS [HR 1.45, 95% credible interval (CrI) 1.16 to 1.82], PFS (HR 1.36, 95% CrI 1.11 to 1.67), overall recurrence [relative risk (RR) 1.19, 95% CrI 1.02 to 1.39] and local recurrence (RR 1.80, 95% CrI 1.19 to 2.71). PAI was worse than RFA for PFS (HR 1.63, 95% CrI 1.05 to 2.51). Resection was better than PEI for OS (HR 0.60, 95% CrI 0.39 to 0.92). RFA combined with PEI decreased the risk of local recurrence compared with PEI alone (RR 0.33, 95% CrI 0.12 to 0.94).

Radiofrequency ablation + iodine-125 appears superior to RFA alone in terms of OS (HR 0.50, 95% CrI 0.31 to 0.80) and overall recurrence (RR 0.69, 95% CrI 0.48 to 0.99). There was also evidence to suggest that RFA + iodine-125 is better than PEI, PAI, TACE + PAI, RFA + TACE and laser ablation for OS, and better than PEI and TACE + PEI for overall recurrence. However, according to our clinical advisors RFA + iodine-125 is only used in selected centres in China.

There was evidence to suggest an increased risk of overall recurrence with MWA + sorafenib, compared with both resection (RR 2.09, 95% CrI 1.12 to 3.89) and RFA + iodine-125 (RR 2.93, 95% CrI 1.31 to 6.56). Also, RFA + systemic chemotherapy decreased the risk of overall recurrence compared with MWA + sorafenib (RR 0.26, 95% CrI 0.08 to 0.92).

The threshold analysis suggested that additional evidence could plausibly change the NMA result for comparisons including RFA, MWA, laser ablation, RFA + TACE, RFA + systemic chemotherapy or RFA + iodine-125. RFA, MWA and laser ablation were agreed to be interventions of interest by the advisory group.

**Systematic review of non-randomised evidence results**

Fourteen non-randomised studies were identified. The majority were conducted in China or Japan, with sample sizes ranging from 21 to 740 patients. No comparative studies were identified on ECT, histotripsy, SABR or wider radiotherapy techniques.

The quality and reporting of the non-randomised studies were poor; 12 had a high RoB. Several studies allocated patients to treatments based on tumour characteristics, so there were potentially prognostic differences between groups at baseline. There was one study with a low RoB. It compared RFA with MWA and included 42 patients. Local tumour progression was similar between groups but new intrahepatic tumours were more frequent in the RFA group. One study of RFA compared with resection had an unclear RoB and included 346 patients. It reported significantly better health-related quality of life (HRQoL), fewer AEs and a shorter hospital stay in the RFA group.

**Updated network meta-analyses and threshold analysis results**

Due to the significant limitations of the non-randomised studies identified, only the two studies that were not at a high RoB were included in the updated NMAs. Additional non-randomised comparative data (RFA vs. MWA vs. IRE) made available prior to publication by a clinical advisor were also included. Updated NMAs using RCT and non-RCT evidence were undertaken for OS, PFS and local recurrence.

Most results of the updated NMAs were consistent with the original results. There remained a high level of uncertainty in treatment rankings. However, the updated NMAs suggested that MWA improves OS and PFS compared with PEI (OS: HR 0.60, 95% CrI 0.40 to 0.90; PFS: HR 0.66, 95% CrI 0.46 to 0.95) and PAI (OS: HR 0.48, 95% CrI 0.24 to 0.99; PFS: HR 0.55, 95% CrI 0.33 to 0.94). Resection also improves PFS compared with PEI (HR 0.72, 95% CrI 0.54 to 0.96) and PAI (HR 0.61, 95% CrI 0.38 to 0.98). The NMA showed IRE to be worse than RFA (RR 2.97, 95% CrI 1.45 to 6.09) and RFA + PEI (RR 4.96, 95% CrI 1.50 to 16.36) for local recurrence, although the CrIs were very wide for both comparisons. There was also evidence that RFA + iodine-125 is better than resection in terms of OS (HR 0.53, 95% CrI 0.30 to 0.94).

The threshold analysis suggested that additional evidence could plausibly change the NMA result for comparisons including MWA, RFA, IRE, RFA + TACE and laser.

**Feasibility of economic modelling**

Limitations in available clinical data may impact the feasibility of undertaking robust economic analysis. However, a value of information (VOI) analysis may be helpful as there are currently several treatments with limited evidence on effectiveness. VOI analysis quantifies the value of reducing decision uncertainty in monetary terms. This can then be compared with the costs of conducting further studies. This could help prioritise which treatments should (or should not) be assessed in future trials. This may be of particular relevance in considering treatments that are currently rarely used in NHS practice but may be effective.

**Patient and public involvement**

The project team included a patient collaborator, who was involved throughout the project. Four additional patients were recruited to the project advisory group, attending meetings at key stages of the project. Patients provided helpful information about the outcomes most important to them, which informed the development of the data extraction tool. Patients were surprised by the lack of data on patient preference and quality-of-life outcomes. Patient and public involvement added context to the review findings and informed the conclusions of the report and recommendations for further research.

**Workshop**

Two workshops were held with clinicians and patients to discuss the project findings and identify key priorities for future research. It was agreed that MWA would be the most appropriate comparator in future trials as it is widely used as the standard of care in the UK, and therapies that are more complex

to deliver were considered unlikely to replace it. MWA is preferred over RFA due to technological advances and ease of use, rather than data on improved clinical effectiveness. However, future research may be most useful if focused on the subgroup of patients with tumours in challenging locations, less fit patients and those with incomplete response to primary therapy. SABR and proton beam therapy were considered to be of particular interest. They are not suitable for patients with advanced or moderately advanced liver disease and, unlike ablation, can usually only be delivered once, but may be appropriate for a subgroup of patients. Histotripsy is at an early stage of regulatory approval, so should not be assessed until efficacy has been demonstrated.

It may be most feasible to undertake an international multicentre RCT as the marginal benefit of novel treatments compared with the existing standard of care is likely to be small, so future studies would need to be large to demonstrate a significant difference in outcomes, and the number of early HCC patients in the UK eligible for all treatments is limited. Outcomes that should be assessed in future trials include local recurrence, overall recurrence, OS, PFS, HRQoL and patient acceptability.

## Conclusions

### *Implications for health care*

There are considerable limitations to the evidence on ablative and non-surgical therapies for early and very early HCC. There is insufficient evidence to draw any conclusions on quality-of-life outcomes. The only firm conclusions that can be drawn from the available data are that PEI and PAI are inferior to RFA, and also appear to be inferior to MWA and resection for certain survival outcomes. MWA and resection are the first-line standard of care for single HCC  $\leq 3$  cm in the UK. The uptake of specific ablative therapies in the UK appears to be based more on technological advancements and ease or speed of use than on high-quality evidence demonstrating superior clinical effectiveness.

### *Recommendations for research*

It is difficult to make firm recommendations for research based on our findings. There are currently no comparative data on several ablative and non-surgical therapies, particularly those treatments reserved for the subgroup of patients with more challenging tumours. However, owing to the small number of such patients who would be eligible for both treatment arms within a trial, along with the marginal benefit of novel treatments compared with the existing standard of care, it is likely to be difficult to recruit sufficient numbers of patients.

Future studies should assess local recurrence, overall recurrence, OS, PFS, HRQoL and patient acceptability, using clear and consistent definitions, in order to allow results to be compared across studies.

Further research on SABR, and possibly other technologies, such as IRE, is required to identify where they should sit in the treatment pathway.

Feasibility studies could address potential issues and complexities in undertaking research in this area prior to undertaking a trial. This would enable: investigation of the acceptability of the intervention (and comparator) to both clinicians and patients, and their willingness to participate in a trial; the practicality of delivering the intervention; and the ability to measure relevant outcomes.

## Study registration

This study is registered as PROSPERO CRD42020221357.

## Funding

This award was funded by the National Institute for Health and Care Research (NIHR) Health Technology Assessment (HTA) programme (NIHR award ref: NIHR131224) and is published in full in *Health Technology Assessment* Vol. 27, No. 29. See the NIHR Funding and Awards website for further award information.

# Chapter 1 Background

Over the last decade, liver cancer incidence has increased by 45% in the UK and is projected to rise further to 15 cases per 100,000 people by 2035.<sup>1</sup> Hepatocellular carcinoma (HCC) is the most common type of primary liver cancer.<sup>2</sup> Between 1997 and 2017 the incidence of HCC in the UK increased by 5.9% a year on average.<sup>3</sup> Primary liver cancer frequently arises on a background of chronic liver disease, and around 90% of cases of HCC are associated with a known underlying aetiology.<sup>2</sup> Globally, hepatitis B virus (HBV) infection is the most common cause of primary liver cancer, but aetiology varies between regions and countries.<sup>4</sup> In the UK, the majority of HCC is associated with the development of cirrhosis, which is most often a consequence of alcohol-related liver disease or non-alcoholic fatty liver disease. Around one-third of patients with cirrhosis develop HCC.<sup>2</sup> Risk increases with the severity of the underlying liver disease in cirrhotic patients,<sup>2</sup> such that patients developing HCC often have advanced liver disease and a significant risk of developing liver failure.

Hepatocellular carcinoma is often asymptomatic until late in its disease course, and the prognosis of HCC patients presenting with symptoms is poor.<sup>5</sup> Recognising the importance of early HCC diagnosis in patients with cirrhosis, the National Institute for Health and Care Excellence (NICE) recommends regular surveillance ultrasound scans intended to diagnose small HCCs so that they can be treated.<sup>6</sup> The Barcelona Clinic Liver Cancer (BCLC) staging system defines very early-stage HCC as a single tumour < 2 cm, preserved liver function and performance status of 0; early-stage disease is defined as a single tumour of any size or up to three tumours ≤ 3 cm, preserved liver function and performance status of 0. Patients with multinodular disease and/or larger tumours would be categorised as having intermediate, advanced or terminal-stage disease (also depending on liver function and performance status).<sup>2</sup> Patients with good liver function who are diagnosed with HCC at an early stage can be offered surgical and non-surgical interventions with curative intent; in general, these patients have favourable 5-year survival rates.<sup>2</sup> However, if patients have signs of advanced cirrhosis with the development of portal hypertension, this restricts the use of liver resection as a treatment option.<sup>7</sup> While liver transplantation is associated with reduced HCC recurrence compared with other treatments, transplantation is limited by availability.<sup>8</sup> Consequently, ablative therapies are frequently used in patients with small HCCs.

A range of ablative and non-surgical therapies is available for treating small HCC tumours in patients with very early or early-stage disease and preserved liver function. The main methods used are microwave ablation (MWA) and radiofrequency ablation (RFA). Alternative methods of ablation include percutaneous ethanol injection (PEI) or percutaneous acetic acid injection (PAI), irreversible electroporation (IRE), laser ablation and cryoablation. Stereotactic ablative radiotherapy [SABR; the term stereotactic body radiotherapy (SBRT) is also used for this technology, but for simplicity SABR is used throughout this report] is emerging as an alternative to invasive ablation and has recently been commissioned as a treatment option by NHS England.<sup>9</sup> Non-ablative approaches, which achieve cure much less frequently, include transarterial (chemo-) embolisation [TA(C)E] and selective internal radiation therapy (SIRT).

However, there has been no definitive assessment of these therapies. NICE guidance comprises overviews of interventional procedures based on rapid reviews, rather than a full systematic assessment of the different treatment options.<sup>10-12</sup> Scoping searches identified four Cochrane Reviews of ablative and minimally invasive therapies that appeared to have populations relevant to this research question; these generally found few or no randomised controlled trials (RCTs), low-quality evidence and a high risk of bias (RoB).<sup>13-16</sup> While some network meta-analyses (NMAs) have been completed, these did not include all relevant therapies and could not assess all relevant outcomes.<sup>17-19</sup> The evidence base is large, but the majority of studies are small and of poor quality. It is also important to consider the applicability of the research evidence to the UK population, since the aetiology of HCC differs between European and Asian populations;<sup>20</sup> many primary studies of interventions for HCC have been undertaken in

## BACKGROUND

Asia. Therefore, a thorough systematic evaluation of the existing research evidence was required to inform UK clinical practice and the design of future effectiveness and cost-effectiveness studies of emerging treatments.

## Chapter 2 Aim and objectives

The aim of this project was to evaluate and compare the effectiveness of ablative and non-surgical therapies for patients with HCC whose tumours are small (up to 3 cm).

The key objectives were:

- to systematically identify all RCTs of ablative and non-surgical therapies for HCC (including registered, unpublished and ongoing trials)
- to evaluate their quality and applicability to UK populations
- to determine the comparative effectiveness of therapies using NMA techniques
- where the evidence base is insufficient, to supplement the RCT evidence with targeted systematic reviews of high-quality, non-randomised, prospective comparative studies of specific therapies
- to identify priority areas where additional high-quality evidence is required (in collaboration with patients and clinicians)
- to assess whether future economic analysis based on the findings would be feasible and worthwhile.



## Chapter 3 Methods

The systematic reviews were conducted following the Centre for Reviews and Dissemination (CRD) guidance on undertaking systematic reviews<sup>21</sup> and reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement.<sup>22,23</sup> The protocol is registered on PROSPERO, the international prospective register of systematic reviews in health and social care; registration number CRD42020221357.

### Systematic review of randomised controlled trials

#### *Search strategy for identification of randomised controlled trials*

A comprehensive, systematic search of bibliographic databases and trial registers was undertaken to identify RCTs of ablative and non-surgical therapies for the treatment of early/small ( $\leq 3$  cm diameter) HCCs. The search strategy was developed in Ovid MEDLINE by an information specialist (MH) with input from the review team. The strategy combined relevant text word searches for terms that appear in the titles or abstracts of database records, with relevant subject headings (e.g. MeSH terms). The strategy consisted of a set of terms for early/small HCC combined with terms for each of the ablative and non-surgical therapies. The MEDLINE search strategy was adapted for use in all other resources searched.

Searches were limited to RCTs using validated study design search filters where available. Retrieval was restricted to articles published from 2000 onwards, as clinical advice confirmed that practice has evolved over the past 20 years and techniques have changed over time. In addition, the natural history and treatment of the underlying liver disease have also changed over the last 20 years, including antiviral therapies for HBV/hepatitis C virus (HCV); therefore, overall outcomes will have changed over this period. Language limits were not applied to the strategy.

The following databases were searched on 3 February 2021:

- MEDLINE ALL (Ovid)
- Embase (Ovid)
- Cochrane Central Register of Controlled Trials (CENTRAL) (Wiley)
- Science Citation Index (Web of Science).

Relevant systematic reviews were also sought, in order to check their reference lists for additional relevant studies. The following systematic review databases were searched on 3 February 2021:

- Cochrane Database of Systematic Reviews (CDSR) (Wiley)
- Database of Abstracts of Reviews of Effects (DARE) (CRD databases)
- International Health Technology Assessment database
- Epistemonikos
- International Prospective Register of Systematic Reviews (PROSPERO).

At our first advisory group meeting on 15 February 2021, a few additional non-surgical therapies were suggested for inclusion in the review: electrochemotherapy (ECT), histotripsy and wider radiotherapy techniques. Therefore, all of the databases listed above were searched again on 17–18 March 2021 using terms for the condition taken from the original searches (devised by MH), with further terms for additional therapies (devised by HF). The records retrieved from these searches were deduplicated against the original search results in EndNote™ 20 (Clarivate Analytics, Philadelphia, PA, USA).

Information on studies in progress and unpublished research was sought by searching ClinicalTrials.gov and the European Union Clinical Trials Register on 27 April 2021, using terms for early/small HCC only. These searches were devised and performed by an information specialist (HF). As trial registers have limited search interfaces which are not designed for expert searches, terms for the condition were searched for without listing any of the interventions, to capture as many relevant records as possible. The search of ClinicalTrials.gov was limited to 'interventional studies', and both registers were limited to trials first posted from 2010 onwards, since the main purpose of searching clinical trial registers was to identify ongoing trials. Clinical advisors were consulted about relevant studies they were aware of.

Search results were imported into EndNote 20 and deduplicated. MEDLINE search strategies are presented in [Appendix 1.1](#). Search strategies for other databases are presented in [Report Supplementary Material 1](#).

### ***Inclusion criteria***

#### **Participants**

Patients diagnosed with HCC with tumour size up to 3 cm (studies with mixed populations were considered if the data for patients with tumour size up to 3 cm could be extracted separately), who were suitable for treatment with ablative or non-surgical therapies. Key participant subgroups considered included:

- size of tumour
- number of tumours (single or multiple lesions)
- disease stage
- cirrhosis and severity (Child–Pugh A or B)
- liver disease (HBV/HCV, other)
- prior HCC treatment
- study location.

#### **Interventions**

Any ablative or non-surgical therapy, including:

- RFA
- MWA
- laser ablation
- high-intensity focused ultrasound (HIFU)
- cryoablation
- PEI
- PAI
- IRE
- TACE
- transarterial embolisation
- SIRT
- ECT
- histotripsy
- SABR
- wider radiotherapy techniques.

#### **Comparators**

The project aimed to evaluate the comparative effectiveness of all of the therapies listed above, so no specific comparator therapy was considered; any comparator was eligible for inclusion, including ablative, minimally invasive or more invasive interventions. Studies comparing a relevant therapy versus surgical resection were also included. Studies comparing different methods of undertaking the same intervention were not eligible for inclusion (e.g. conventional temperature control RFA vs. impedance

control RFA, RFA under ultrasound guidance vs. RFA under computed tomography guidance); studies had to compare two different therapies.

## Outcomes

The outcomes of interest were:

- overall survival (OS)
- progression-free survival (PFS)
- time to progression (TTP)
- recurrence
- serious adverse events (AEs)
- intervention-specific AEs (e.g. pneumothorax, post-ablation syndrome, post-embolisation syndrome, thermoablative injury, pain, haemorrhage or bile leak)
- quality of life.

Where reported, outcomes of economic relevance were recorded, including healthcare costs and duration of hospital stay.

## Study design

Randomised controlled trials were eligible for inclusion.

### Study selection and data extraction

Studies were initially assessed for relevance using titles and abstracts. As the database searches were expected to be extensive, a single reviewer screened each identified title/abstract, and 10% of records were checked by another reviewer. Full-text articles were independently screened by two reviewers for final inclusion. Any disagreements were resolved through discussion and, where necessary, consultation with a third reviewer. Foreign-language studies were translated and assessed for inclusion. Studies only available as conference abstracts were identified and attempts were made to contact authors for further data to enable them to be assessed for inclusion in the review.

A data extraction form was developed using Microsoft Excel® (Microsoft Corporation, Redmond, WA, USA), piloted on a sample of studies and refined. Data on intervention, comparator and patient characteristics and results were extracted by one reviewer (SS-H or ES) and checked by a second reviewer (RW). Any discrepancies were resolved by discussion. Foreign-language studies were data extracted by a native speaker and discussed at a meeting with a second reviewer (RW). Authors of conference abstracts were contacted for further information; data were extracted using only the abstract when authors did not respond.

For all outcomes, data were extracted from publications either as hazard ratios (HRs) for survival outcomes, or as relative risks (RRs) for dichotomous outcomes, and in all cases with their corresponding 95% confidence intervals (CIs) or standard errors (SEs).

For survival outcomes, where studies did not report HRs and their variances, Kaplan–Meier (KM) data, including the numbers at risk, were extracted using methods reported by Guyot *et al.*<sup>24</sup> and HRs were computed using the reconstructed individual patient data. If a study did not report the numbers at risk, the p-value for the log-rank test was used to calculate the HR and its corresponding variance using methods described by Irvine *et al.*<sup>25</sup>

In the instance where neither HRs were reported nor KM plots were provided, HRs and SEs were back-calculated using the reported survival rates and the p-value of the log-rank test with the log-rank test.<sup>26</sup>

### Critical appraisal

Risk of bias in RCTs was assessed using the latest version of the Cochrane RoB tool.<sup>27</sup> RoB assessment was undertaken by one reviewer (SS-H or ES) and independently checked by a second reviewer (RW). Any disagreements were resolved through consensus.

## Network meta-analysis

### Feasibility assessment

Randomised controlled trials were mapped according to interventions included, outcomes reported, trial size and quality, to determine the overall extent of the RCT evidence. Trials were grouped according to identified subgroups (e.g. tumour size and stage), where appropriate. Key interventions and comparisons of interventions where existing RCT data are absent, limited or of poor quality were identified. The mapping was used to determine whether NMA of the RCTs was feasible.

Networks of treatment comparisons were drawn for each outcome to check that they were connected. Not all RCTs reported data that could be used; only studies with usable data were included in the networks.

### Included data

Network meta-analyses were conducted for four outcomes: OS, PFS, overall recurrence, and local recurrence. For OS and PFS, only contrast-level data were available in the form of HRs. For overall recurrence and local recurrence, both contrast-level and arm-level data were available. Data for both HRs and RRs were synthesised on the log scale, by log-transforming estimates and their CIs from studies.

For OS and PFS, summary effect estimates from the NMAs were presented as HRs and their corresponding 95% credible intervals (CrIs), whereas overall and local recurrence estimates were presented as RRs and their corresponding 95% CrIs.

Any deviation from proportional hazards was tested for, and the Schoenfeld residuals, survival curves and piecewise hazards visually inspected. If there is strong evidence that the proportional hazards assumption does not hold, or the simpler models initially considered do not fit the data well, more complex, time-varying models that account for non-proportional hazards should be considered, if sufficient data are available. However, data were limited, so this was not possible. Consequently, appropriate caution with the results is expressed, where appropriate.

### Network meta-analysis

Network meta-analyses were conducted in a Bayesian framework using Markov chain Monte Carlo techniques. For the aggregate RCT data (HRs and RRs), contrast-based models proposed by Dias *et al.*, which appropriately account for correlations in trials with more than two arms, were used.<sup>28–30</sup> All four outcomes were modelled using a normal likelihood with an identity link.<sup>30</sup> Where arm-level data were available for overall and local recurrence, the binomial likelihood, logit link model suggested by Warn *et al.*<sup>31</sup> was also fitted to prove comparability of the results.

All analyses were carried out using the GeMTC package<sup>32</sup> in R (version 4.1.2).

To account for the correlation between the relative effects in three-arm trials<sup>33</sup> the covariance between differences taken with respect to the same control arm was calculated using the equation:

$$\text{Cov}(y_{ab}, y_{ac}) = \frac{\text{Var}(y_{ab}) + \text{Var}(y_{ac}) - \text{Var}(y_{bc})}{2} \quad (1)$$

Fixed-effect (FE) and random-effects (RE) models were fitted. Models were sampled for 100,000 iterations over four chains after an initial burn-in of 50,000 iterations. Model convergence was assessed through visual inspection of Brook–Gelman–Rubin diagnostic and history plots.<sup>34</sup>

For the RE models, the choice of prior distributions for the between-study standard deviation (SD) was explored. A half-normal (0, 0.19<sup>2</sup>) and a uniform (0, 3) prior distribution were considered. As a sensitivity analysis, a half-normal (0, 0.50<sup>2</sup>) prior was also used for the between-study heterogeneity.<sup>35</sup>

Models were compared based on their deviance information criteria (DIC), and the model with the smallest DIC was selected as the base-case analysis.<sup>36,37</sup> Differences < 3 were not considered meaningful, and the simplest model was selected. Where a FE model was selected, results for the RE models were also presented as a sensitivity analysis.

In networks with loops formed by independent studies (i.e. where different studies provided direct and indirect evidence for the same comparison), inconsistency (i.e. conflict between the direct and indirect evidence) was checked by comparing the model fit and between-study heterogeneity from the NMA models versus the corresponding unrelated mean effects (inconsistency) models.<sup>28,38</sup> Where inconsistency was identified, it was explored by inspecting the characteristics of the included studies (participant and design characteristics) that may contribute to inconsistency. Where feasible, node-split models were fitted to provide further evidence of the location and impact of potential inconsistency.<sup>39</sup>

Where judged appropriate, NMA was used to assess and rank interventions by comparative effectiveness. Where feasible, the potential impact of additional evidence on the NMAs was investigated using threshold analysis.<sup>40,41</sup>

## Threshold analysis

Threshold analysis<sup>40,41</sup> was conducted at the contrast level to examine the impact of potential changes to the evidence on each treatment contrast to identify which treatment comparisons lacked robust RCT evidence. Threshold analysis represents a robust statistical alternative to qualitative assessment of the robustness of evidence. It is a novel statistical approach that can be used to investigate which comparisons in a NMA have estimated relative effects which might not be robust to changes in the observed evidence due to either possible bias, sampling variation or relevance.<sup>40,41</sup> Threshold analysis uses formal statistical methods to quantify precisely how much the results of a NMA could vary (due to changes in the amount of data, or due to potential bias) before any conclusion changes (e.g. changes to the ranking of an intervention), by examining what the smallest changes to the available data required to alter a conclusion are. It can therefore be used to identify which interventions, or comparisons of interventions, have the most robust evidence, and which interventions would benefit from further trials.

Threshold analysis was carried out using the *nmathresh* package<sup>40</sup> in R (version 4.1.2). Results of the threshold analysis are presented graphically as forest plots and threshold tables. The results have been used to identify interventions and comparisons where non-randomised evidence should be sought for further review, based on the sensitivity shown by the comparison with potential additional evidence.

Following clinical advice, comparisons that included PAI and PEI were excluded from the threshold analysis to restrict attention to interventions considered relevant to current practice.

## Systematic review of non-randomised evidence

Results of the mapping exercise, NMAs and threshold analyses were used to identify interventions or comparisons where non-randomised evidence might usefully add to the RCT evidence or potentially

resolve uncertainty (see [Systematic review of RCTs](#), [Network meta-analysis results](#) and [Threshold analysis of RCT networks](#)). This identified and classified evidence for interventions:

1. with no RCT evidence
2. with limited RCT evidence (e.g. only one or two trials, or <50 or <100 patients in total)
3. where RCT evidence is very heterogeneous (e.g. very different results across trials)
4. where RCT evidence is highly uncertain (e.g. wide CIs or uncertain ranking in NMAs, as identified by the threshold analysis)
5. where RCT evidence is of low or uncertain quality, or at ROB.

The advisory group was consulted to identify interventions of particular practical interest where RCT evidence was lacking. A distinction was made between comparisons without any current RCT evidence (i.e. where an intervention of interest was disconnected from the main network) and comparisons with imprecise or non-robust RCT evidence.

This targeted approach was used because preliminary searches suggested that the quantity of non-randomised evidence was too large to be fully reviewed within the time and resource available for this project; furthermore, this would be of limited value as much of the non-randomised evidence is likely to be of insufficient quality for inclusion in any analysis.

For the interventions identified for further investigation by our classification and by the advisory group, targeted database searching and screening were performed.

### **Search strategy for identification of non-RCTs**

Searches were undertaken to identify non-randomised studies of selected interventions for early/small ( $\leq 3$  cm diameter) HCC, where RCT evidence was not available. The search strategy consisted of terms for small or early HCC combined with terms for the selected interventions (HIFU, cryoablation, IRE, ECT, histotripsy, SABR and wider radiotherapy techniques). Relevant subject headings alongside text word searches in the title and abstracts of records were included in the search strategy. To allow comprehensive retrieval of non-randomised studies, the search was not restricted by study type.<sup>42</sup> The strategy was limited to articles published from the year 2000 onwards. Language limits were not applied.

The searches were carried out on 28 July 2021. The following databases were searched: MEDLINE (Ovid), Embase (Ovid), CENTRAL (Wiley) and the Science Citation Index (Web of Science, Clarivate). EndNote 20 was used to manage and deduplicate the search results.

Although conference abstracts were due to be identified via a search of the Conference Proceedings Citation Index – Science, a pragmatic decision to not search this database was taken due to a lack of time and resources to screen and follow up ongoing studies reported as conference abstracts. Similarly, conference abstracts were removed from the search results retrieved in Embase.

MEDLINE search strategies are presented in [Appendix 1.2](#). Search strategies for other databases are presented in [Report Supplementary Material 1](#).

Searches were also undertaken to identify studies of selected interventions for comparisons where additional evidence could plausibly change the NMA conclusions, as identified by the threshold analysis. The search strategy consisted of terms for small or early HCC combined with terms for the selected interventions (RFA, MWA and laser ablation, compared with each other or with surgical resection). Relevant subject headings alongside text word searches in the title and abstracts of records were included in the search strategy. The strategy was limited to articles published from the year 2000 onwards, and animal studies were removed where possible. Language limits were not applied.

The searches were carried out on 24 August 2021. The following databases were searched: MEDLINE (Ovid), Embase (Ovid), CENTRAL (Wiley) and the Science Citation Index (Web of Science, Clarivate). EndNote 20 was used to manage and deduplicate the search results.

MEDLINE search strategies are presented in [Appendix 1.3](#). Search strategies for other databases are presented in [Report Supplementary Material 1](#).

## **Inclusion criteria**

### **Participants**

Patients diagnosed with HCC with tumour size up to 3 cm (studies with mixed populations were considered if the data for patients with tumour size up to 3 cm could be extracted separately), who were suitable for treatment with ablative or non-surgical therapies. Studies of patients with recurrent HCC were excluded, as clinical advisors confirmed that it was not appropriate to synthesise the results of these studies with the studies of HCC patients included in the networks.

### **Interventions**

Informed by the systematic review of RCTs and results of the NMAs and threshold analyses (see [Systematic review of RCTs](#), [Network meta-analysis results](#) and [Threshold analysis of RCT networks](#)), ablative or non-surgical therapies of particular practical interest where RCT evidence was lacking were sought; these were interventions where either RCT evidence was not available, or where additional evidence could plausibly change the NMA result, as identified by the threshold analysis. The specific interventions were:

- RFA
- MWA
- laser ablation
- HIFU
- cryoablation
- IRE
- ECT
- histotripsy
- SABR
- wider radiotherapy techniques.

### **Comparators**

The project aimed to evaluate the comparative effectiveness of all the therapies listed above, so no specific comparator therapy was considered; any comparator was eligible for inclusion, including ablative, minimally invasive or more invasive interventions. Studies comparing a relevant therapy versus surgical resection were also included. Studies comparing different methods of undertaking the same intervention were not eligible for inclusion (e.g. conventional temperature control RFA vs. impedance control RFA; RFA under ultrasound guidance vs. RFA under computed tomography guidance); studies had to compare two different therapies.

### **Outcomes**

The outcomes of interest were:

- OS
- PFS
- TTP
- quality of life.

Studies only reporting response and AE results were excluded from the review of non-RCTs as these outcomes were not relevant for the NMAs.

### Study design

Only prospective non-randomised studies that compared two or more eligible therapies were included; studies of single therapies were excluded.

### Study selection and data extraction

Consistent with the review of RCTs, titles and abstracts were screened by a single reviewer, with 10% of records checked by another reviewer. Full-text articles were independently screened by two reviewers for final inclusion. Any disagreements were resolved through discussion and, where necessary, consultation with a third reviewer. Foreign-language studies were translated and assessed for inclusion. Studies only available as conference abstracts were assessed based on the limited data available and were included if there were sufficient data reported on the relevant outcomes.

The data extraction form developed using Microsoft Excel for the review of RCTs was modified for the review of non-RCTs. Data on intervention, comparator and patient characteristics and results were extracted by one reviewer (RW or ES) and independently checked by a second reviewer (ES or RW). Any discrepancies were resolved by discussion. Foreign-language studies were data extracted by a native speaker and discussed at a meeting with a second reviewer (RW). Where studies were only reported as conference abstracts, data were extracted using the limited data available. Where possible, HRs and their variances were extracted by one reviewer and checked by a second reviewer. When the HRs were not available, KM data were extracted using methods reported by Guyot *et al.*<sup>24</sup> If neither HRs nor KM data were available, survival rates and p-values for the log-rank test were extracted.

### Critical appraisal

A validity assessment tool was developed, piloted on a sample of studies and refined. Validity assessment was undertaken by one reviewer (RW or ES) and independently checked by a second reviewer (ES or RW). Any disagreements were resolved through consensus. The most important quality assessment criteria were selected, based on their potential impact on the overall validity of the studies, and an overall RoB judgement was made for each study; important criteria were those relating to the participant inclusion criteria, appropriateness of treatment allocation, similarity of treatment groups at baseline and whether missing outcome data were balanced across treatment groups.

### Updated network meta-analysis

For non-randomised studies that were of sufficient quality, the NMA and threshold analyses were repeated after pooling (without any adjustments) the non-randomised evidence with the RCT evidence, to assess whether estimates were improved.

The updated NMA was conducted using the methods detailed in *Network meta-analysis*.

### Updated threshold analysis

A threshold analysis was conducted on the results for the updated NMAs using both RCT and non-randomised evidence to explore the robustness of the updated results.

The updated threshold analysis was conducted using methods detailed in *Threshold analysis*.

# Chapter 4 Results

## Systematic review of RCTs

The electronic searches identified a total of 7550 records after deduplication between databases; 6774 records were identified from the original searches of bibliographic databases undertaken on 3–4 February 2021, 655 from the searches for studies of ECT, histotripsy and wider radiotherapy techniques undertaken on 17–18 March 2021, and 121 from the trial register searches undertaken on 27 April 2021. One additional record was identified from screening reference lists of relevant systematic reviews. Clinical advisors were not aware of any additional RCTs not identified in the electronic searches.

Two hundred potentially relevant studies were ordered for full paper screening. Twenty-seven full papers were unavailable as they were only reported as conference abstracts or clinical trial register records; study authors were e-mailed (where contact details could be found) and authors of six records confirmed that they were either duplicate reports or did not meet our inclusion criteria. One hundred and seventy-three full papers were screened; 138 were excluded at the full paper stage and are listed in [Appendix 2](#), along with the reasons for their exclusion. [Figure 1](#) presents the flow of studies through the study selection process.

### Characteristics of RCTs included in the review

Details of the 37 RCTs that were included in the systematic review are presented in [Table 1](#). One RCT was ongoing and therefore no results were available for data extraction. The characteristics and results of the other 36 RCTs were extracted into an Excel spreadsheet.

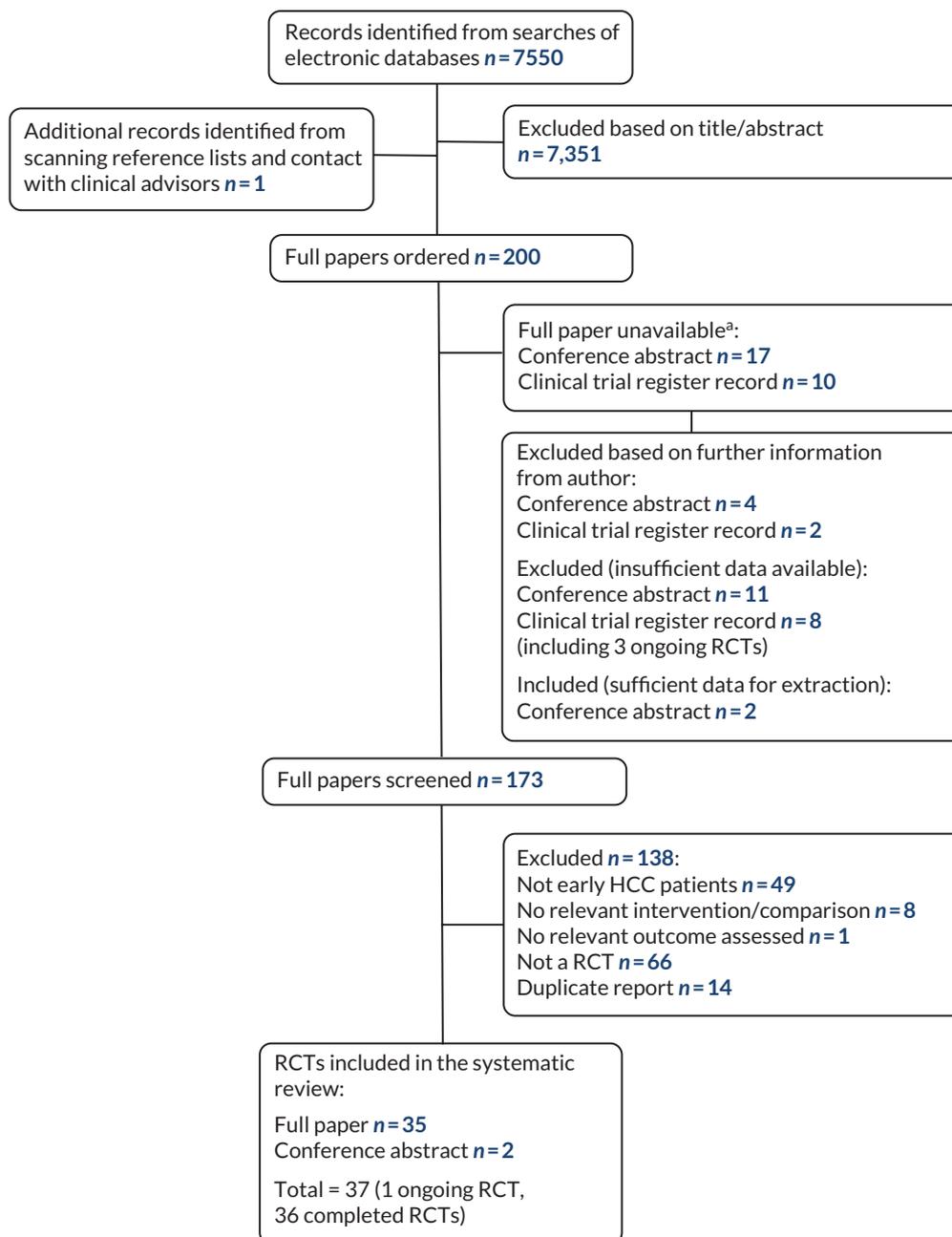
Fifteen of the 36 completed RCTs restricted inclusion criteria to HCC patients with tumour size up to 3 cm in diameter.<sup>43–57</sup> Six RCTs included patients with tumours up to 4 cm in diameter,<sup>58–63</sup> 12 RCTs included patients with tumours up to 5 cm in diameter<sup>64–75</sup> and one RCT included patients with tumours up to 7 cm in diameter.<sup>76</sup> One RCT did not report specific tumour size criteria but included patients with small HCCs,<sup>77</sup> and one RCT included patients within BCLC stages 0–B.<sup>78</sup> The RCTs that included patients with larger tumours (>3 cm diameter) were included in the review if they reported separate results for the subgroup of patients with a tumour diameter up to 3 cm or, in the case of three RCTs, if a clear majority of patients had tumours < 3 cm in diameter.<sup>60,63,71</sup> Three RCTs included patients with recurrent/residual tumours ≤ 3 cm.<sup>51,74,75</sup> Sample sizes ranged from 30 to 308 patients.

The majority of RCTs were conducted in Asian countries, which has implications for the generalisability of results to the UK population. HCC in European patients is more likely to be caused by alcohol or hepatitis C, whereas in Asia it is more likely to be caused by hepatitis B. The natural history of these diseases is different and treatment options for the underlying liver disease differ. RCTs were conducted in China ( $n = 17$ ), Japan ( $n = 7$ ), Taiwan ( $n = 4$ ), South Korea ( $n = 1$ ), Egypt ( $n = 2$ ), Italy ( $n = 4$ ), Italy and Germany ( $n = 1$ ) and Switzerland and France ( $n = 1$ ).

The most frequently assessed ablative/non-surgical therapy was RFA, either alone or in combination with TACE, PEI, iodine-131 metuximab, iodine-125 or chemotherapy. [Table 2](#) shows the comparisons made in the included RCTs. The majority of RCTs assessed OS, progression-/disease-free survival and/or recurrence, along with response and AEs. A few RCTs presented economic outcomes. Only one RCT assessed patient satisfaction.

### Quality of RCTs included in the review

Risk of bias was assessed for each of the main study outcomes using the Cochrane RoB tool, resulting in 58 assessments for the 35 included RCTs for which RoB could be assessed; two RCTs did not have a RoB assessment as they were either ongoing<sup>80</sup> or did not report any relevant outcomes for the subgroup of



**FIGURE 1** Flow diagram of the study selection process (RCTs).

<sup>a</sup>Where possible, authors were contacted for further information.

patients with tumours  $\leq 3$  cm.<sup>58</sup> Results of the RoB assessment for the most relevant outcome assessed are presented in [Table 3](#). Results for each of the main study outcomes are presented in [Appendix 3](#). Two RCTs were only reported as conference abstracts; therefore, some questions had a 'no information' response owing to the limited reporting, resulting in a high RoB for the domain and the study overall.<sup>43,50</sup>

Generally, methods were poorly reported. There was either a high RoB or some concerns arising from the randomisation process in 20/35 of the RCTs assessed. Most RCTs had a low RoB for domains relating to deviations from the intended intervention (27/35), missing outcome data (24/35) and selective outcome reporting (34/35 had a low RoB for the most relevant outcome). All RCTs had a low RoB relating to measurement of the outcome, using computerised tomography (CT) (or magnetic

**TABLE 1** RCTs included in the systematic review

Study	Location	Participant information	Intervention	Comparator
Abdelaziz, 2014 <sup>64</sup>	Egypt	111 patients (with 128 tumours) $\leq$ 5 cm; subgroup of 87 tumours $\leq$ 3 cm	RFA	MWA
Aikata, 2006 (abstract) <sup>43</sup>	Not reported (authors from Japan)	44 patients with tumours $<$ 3 cm	RFA + TACE	RFA alone
Azab, 2011 <sup>65</sup>	Egypt	90 patients (with 98 tumours) $\leq$ 5 cm; subgroup of 48 tumours $\leq$ 3 cm	PEI + RFA	RFA alone PEI alone
Bian, 2014 <sup>78</sup>	China	127 patients with BCLC stage 0–B; subgroup of 78 patients with tumours $<$ 3 cm	RFA + iodine-131 metuximab	RFA alone
Brunello, 2008 <sup>44</sup>	Italy	139 patients with tumours $\leq$ 3 cm	RFA	PEI
Chen, 2005 (reported in Chinese) <sup>66</sup>	China	86 patients with tumours $\leq$ 5 cm; subgroup of 47 patients with tumours $\leq$ 3 cm	RFA + PEI	RFA alone
Chen, 2005 (reported in Chinese) <sup>67</sup>	China	132 patients with tumours $\leq$ 5 cm; subgroup of 55 patients with tumours $\leq$ 3 cm	Resection	RFA
Chen, 2006 <sup>68</sup>	China	180 patients with tumours $\leq$ 5 cm; subgroup of 79 patients with tumours $\leq$ 3 cm	Percutaneous local ablative therapy	Partial hepatectomy
Chen, 2014 <sup>45</sup>	China	136 patients with tumours $\leq$ 3 cm	RFA + iodine-125	RFA alone
Fang, 2014 <sup>46</sup>	China	120 patients with tumours $\leq$ 3 cm	RFA	Hepatectomy
Feng, 2012 <sup>58</sup>	China	168 patients with tumours $<$ 4 cm; subgroup of 56 patients with tumours $\leq$ 2 cm	RFA	Surgical resection
Ferrari, 2007 <sup>77</sup>	Not reported (authors from Italy)	81 patients with tumours $\leq$ 4 cm; subgroup of 28 patients with tumours $\leq$ 2.5 cm	Laser ablation	RFA
Gan, 2004 (reported in Chinese) <sup>47</sup>	China	38 patients with tumours $\leq$ 3 cm	RFA alone	RFA + chemotherapy
Giorgio, 2011 <sup>48</sup>	Italy	285 patients with tumours $\leq$ 3 cm	RFA	PEI
Huang, 2005 <sup>49</sup>	Taiwan	82 patients with tumours $\leq$ 3 cm	PEI	Resection
Huang, 2010 <sup>69</sup>	China	230 patients with tumours $\leq$ 5 cm; subgroup of 159 patients with tumours $\leq$ 3 cm	RFA	Resection
Huo, 2003 <sup>70</sup>	Taiwan	108 patients with tumours $\leq$ 5 cm; subgroup of 55 patients with tumours $\leq$ 3 cm	Sequential TACE and PAI	PAI alone
Izumi, 2019 (abstract) <sup>50</sup>	Japan	308 patients with tumours $\leq$ 3 cm	RFA	Surgery
Kim, 2020 <sup>51</sup>	South Korea	144 patients with recurrent/residual tumours $<$ 3 cm	Proton beam radiotherapy	RFA
Koda, 2001 <sup>52</sup>	Japan	52 patients with tumours $<$ 3 cm	TACE + PEI	PEI alone
Lencioni, 2003 <sup>71</sup>	Not reported (authors from Italy and Germany)	104 patients with tumours $\leq$ 5 cm (large proportion had tumours $\leq$ 3 cm)	PEI	RFA
Lin, 2004 <sup>59</sup>	Not reported (authors from Taiwan)	157 patients with tumours $\leq$ 4 cm; subgroup of 114 patients with tumours $\leq$ 3 cm	RFA	Low-dose PEI High-dose PEI

continued

**TABLE 1** RCTs included in the systematic review (*continued*)

Study	Location	Participant information	Intervention	Comparator
Lin, 2005 <sup>53</sup>	Taiwan	187 patients with tumours $\leq$ 3 cm	RFA	PEI PAI
Liu, 2016 <sup>72</sup>	China	200 patients with tumours $\leq$ 5 cm; subgroup of 135 patients with tumours $\leq$ 3 cm	Partial hepatectomy	TACE + RFA
Mizuki, 2010 <sup>60</sup>	Japan	30 patients with tumours $\leq$ 4 cm (large proportion had tumours $\leq$ 3 cm)	PEI alone	TACE + PEI
Ng, 2017 <sup>79</sup>	China	218 patients with tumours $\leq$ 5 cm; subgroup of 55 patients with tumours $\leq$ 2 cm	Resection	RFA
Orlacchio, 2014 <sup>61</sup>	Italy	30 patients with tumours $\leq$ 4 cm (mean tumour size 2.4 cm)	Laser ablation	RFA
Peng, 2012 <sup>74</sup>	China	139 patients with recurrent HCC tumours $\leq$ 5 cm; subgroup of 87 patients with tumours $\leq$ 3 cm	RFA + TACE	RFA alone
Shibata, 2002 <sup>62</sup>	Japan	72 patients (with 94 tumours) $<$ 4 cm; subgroup of 88 tumours $\leq$ 3 cm	RFA	MWA
Shibata, 2009 <sup>54</sup>	Japan	89 patients with tumours $\leq$ 3 cm	RFA + TACE	RFA alone
Shiina, 2005 <sup>55</sup>	Japan	232 patients with tumours $\leq$ 3 cm	RFA	PEI
Vietti Violi, 2018 <sup>63</sup>	Switzerland and France	152 patients with tumours $\leq$ 4 cm (mean tumour size 1.8 cm, $<$ 8% patients had tumours $>$ 3 cm)	MWA	RFA
Xia, 2020 <sup>75</sup>	China	240 patients with recurrent HCC tumours $\leq$ 5 cm; subgroup of 159 patients with tumours $\leq$ 3 cm	RFA	Repeat hepatectomy
Yan, 2016 <sup>56</sup>	China	120 patients with tumours $\leq$ 3 cm	Resection	MWA + sorafenib
Zhang, 2007 <sup>76</sup>	China	133 patients with tumours $\leq$ 7 cm; subgroup of 60 patients with tumours $\leq$ 3 cm	RFA + PEI	RFA alone
Zhu, 2021 (protocol) <sup>80</sup>	China	Ongoing RCT	RFA	Laparoscopic hepatectomy
Zou, 2017 <sup>57</sup>	China	74 patients with tumours $\leq$ 3 cm	Laser ablation	RFA

resonance imaging) for assessment of tumour response, progression and recurrence. The overall judgement of RoB was low for 9 RCTs and high for 12 RCTs, and there were some concerns for 14 RCTs.

### Results of RCTs included in the review

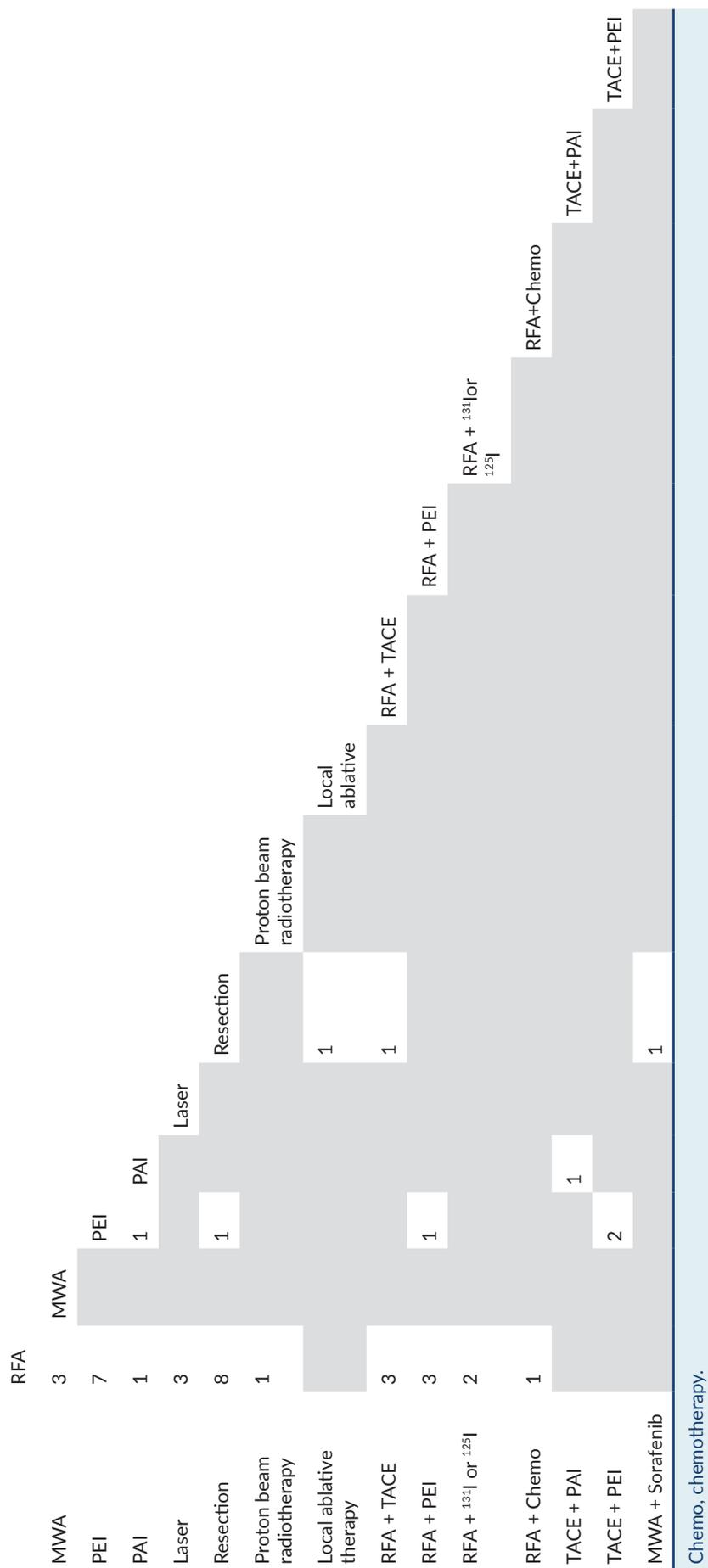
A table of study characteristics and results is presented in [Appendix 4](#).

### Radiofrequency ablation versus microwave ablation

Three RCTs compared RFA with MWA. One was assessed as having a high RoB<sup>64</sup> and the other two as having some concerns.<sup>62,63</sup> One RCT included 152 participants with tumours up to 4 cm but only a small minority of patients had tumours  $>$  3 cm.<sup>63</sup> The other two RCTs only reported the number of tumours  $\leq$  3 cm ( $n = 87$  and  $n = 88$ ) rather than the number of patients.<sup>62,64</sup>

Only one RCT (with some RoB concerns) reported OS and recurrence outcomes.<sup>63</sup> OS was similar between the two treatment groups at 2 years (RFA 84% vs. MWA 86%). More patients in the RFA group had experienced recurrence (local tumour progression) at 2 years (12% vs. 6%; RR 1.62, 95% CI 0.66 to

**TABLE 2** Number of RCTs making each comparison



RESULTS

**TABLE 3** Risk of bias assessment results (RCTs)

Trial	ROB arising from the randomisation process	ROB due to deviations from the intended intervention	ROB due to missing outcome data	ROB in measurement of the outcome	ROB in selection of the reported result	Overall judgement of ROB
Abdelaziz, 2014 <sup>64</sup>	High	Low	Low	Low	Low	High
Aikata, 2006 (abstract) <sup>43</sup>	Some concerns	High	High	Low	Low	High
Azab, 2011 <sup>65</sup>	Some concerns	Low	Low	Low	Low	Some concerns
Bian, 2014 <sup>78</sup>	Some concerns	Low	Low	Low	Low	Some concerns
Brunello, 2008 <sup>44</sup>	Low	Low	Low	Low	Low	Low
Chen, 2005 <sup>67</sup>	Some concerns	Low	Low	Low	Low	Some concerns
Chen, 2005 <sup>66</sup>	Some concerns	Low	High	Low	Low	High
Chen, 2006 <sup>68</sup>	Low	High	Low	Low	Low	High
Chen, 2014 <sup>45</sup>	Low	Low	Low	Low	Low	Low
Fang, 2014 <sup>46</sup>	Some concerns	Low	Low	Low	Low	Some concerns
Ferrari, 2007 <sup>77</sup>	Some concerns	Low	Low	Low	Low	Some concerns
Gan, 2004 <sup>47</sup>	Some concerns	Low	High	Low	Low	High
Giorgio, 2011 <sup>48</sup>	Low	High	High	Low	Low	High
Huang, 2005 <sup>49</sup>	High	High	High	Low	Low	High
Huang, 2010 <sup>69</sup>	Low	Low	High	Low	Low	High
Huo, 2003 <sup>70</sup>	High	Low	High	Low	Low	High
Izumi, 2019 (abstract) <sup>50</sup>	Some concerns	High	High	Low	Some concerns	High
Kim, 2020 <sup>51</sup>	Some concerns	Low	Low	Low	Low	Some concerns
Koda, 2001 <sup>52</sup>	Low	Low	Low	Low	Low	Low
Lencioni, 2003 <sup>71</sup>	Low	Low	Low	Low	Low	Low
Lin, 2004 <sup>59</sup>	Low	Low	Low	Low	Low	Low
Lin, 2005 <sup>53</sup>	Some concerns	Low	Low	Low	Low	Some concerns
Liu, 2016 <sup>72</sup>	Some concerns	Low	Low	Low	Low	Some concerns
Mizuki, 2010 <sup>60</sup>	Low	Some concerns	Some concerns	Low	Low	Some concerns
Ng, 2017 <sup>79</sup>	Low	Low	Low	Low	Low	Low
Orlacchio, 2014 <sup>61</sup>	Some concerns	Low	Some concerns	Low	Low	Some concerns

**TABLE 3** Risk of bias assessment results (RCTs) (continued)

Trial	ROB arising from the randomisation process	ROB due to deviations from the intended intervention	ROB due to missing outcome data	ROB in measurement of the outcome	ROB in selection of the reported result	Overall judgement of ROB
Peng, 2012 <sup>74</sup>	Low	Low	Low	Low	Low	Low
Shibata, 2002 <sup>62</sup>	Low	Low	Some concerns	Low	Low	Some concerns
Shibata, 2009 <sup>54</sup>	High	Low	Low	Low	Low	High
Shiina, 2005 <sup>55</sup>	Some concerns	Low	Low	Low	Low	Some concerns
Vietti Violi, 2018 <sup>63</sup>	Low	Some concerns	Low	Low	Low	Some concerns
Xia, 2020 <sup>75</sup>	Low	Low	Low	Low	Low	Low
Yan, 2016 <sup>56</sup>	High	Some concerns	Low	Low	Low	High
Zhang, 2007 <sup>76</sup>	Low	Low	Low	Low	Low	Low
Zou, 2017 <sup>57</sup>	Some concerns	Low	Low	Low	Low	Some concerns
<b>Total</b>	<b>High: 5 Some concerns: 15 Low: 15</b>	<b>High: 5 Some concerns: 3 Low: 27</b>	<b>High: 8 Some concerns: 3 Low: 24</b>	<b>Low: 35</b>	<b>Some concerns: 1 Low: 34</b>	<b>High: 12 Some concerns: 14 Low: 9</b>

3.94), but the median TTP was longer after RFA than after MWA (16 months vs. 12 months; HR 0.72, 95% CI 0.44 to 1.18).

There was a high rate of complete response or complete ablation of tumours in both the RFA and MWA arms in all three RCTs. A slightly higher proportion of HCC nodules showed complete response after RFA in one RCT (96% vs. 89%),<sup>62</sup> whereas in the other two RCTs the rates were similar between treatment arms.

One RCT reported a higher rate of major complications with MWA than with RFA (RFA 3% vs. MWA 11%).<sup>62</sup> Another RCT reported that grade IV AEs only occurred in the MWA arm (0 vs. 2%), but more grade III (3% vs. 0%) and grade I–II (11.5% vs. 5%) AEs occurred in the RFA arm.<sup>63</sup> The RCT at high RoB reported that there were no major complications in either group.<sup>64</sup>

### Radiofrequency ablation versus percutaneous ethanol injection

Seven RCTs compared RFA with PEI ( $n = 1061$  patients in six RCTs; the other RCT included 48 tumours). Three RCTs had a low RoB,<sup>44,59,71</sup> three were judged to have some concerns<sup>53,55,65</sup> and one had a high RoB.<sup>48</sup> One RCT included two different PEI arms with either a low dose or a high dose of PEI.<sup>59</sup> One RCT compared RFA versus PEI versus RFA in combination with PEI; the results of the combined RFA + PEI group are reported in the relevant sections below.<sup>65</sup> One RCT included patients with tumours  $\leq 5$  cm, but a large proportion had tumours  $\leq 3$  cm.<sup>71</sup>

Six of the seven RCTs reported OS (see [Table 4](#)).<sup>44,48,53,55,59,71</sup> OS was better after treatment with RFA in four of the RCTs, which were at low RoB<sup>59,71</sup> or had some concerns.<sup>53,55</sup> OS was similar between groups in one high-quality RCT<sup>44</sup> and one low-quality RCT.<sup>48</sup>

**TABLE 4** Radiofrequency ablation vs. PEI – OS

	RFA	PEI	High-dose PEI	Study
1 year	95%	95%	-	Giorgio, 2011
	100%	96%	-	Lencioni, 2003
	1–2 cm: 96% / 2.1–3 cm: 89%	1–2 cm: 94% / 2.1–3 cm: 84%	1–2 cm: 93% / 2.1–3 cm: 83%	Lin, 2004
	93%	88%	-	Lin, 2005
2 years	90%	83%	-	Giorgio, 2011
	98%	88%	-	Lencioni, 2003
	1–2 cm: 84% / 2.1–3 cm: 78%	1–2 cm: 78% / 2.1–3 cm: 70%	1–2 cm: 80% / 2.1–3 cm: 71%	Lin, 2004
	81%	66%	-	Lin, 2005
3 years	26 deaths/70 patients	28 deaths/69 patients	-	Brunello, 2008
	83%	78%	-	Giorgio, 2011
	1–2 cm: 78% / 2.1–3 cm: 73%	1–2 cm: 70% / 2.1–3 cm: 62%	1–2 cm: 72% / 2.1–3 cm: 64%	Lin, 2004
	74%	51%	-	Lin, 2005
4 years	73%	70%	-	Giorgio, 2011
	74%	57%	-	Shiina, 2005
5 years	70%	68%	-	Giorgio, 2011

Event-free survival (survival free of local recurrence, new HCC and extrahepatic metastases) was also higher after RFA than after PEI in one high-quality RCT (2-year rate: 64% vs. 43%).<sup>71</sup> Two RCTs (one high quality,<sup>59</sup> one with some concerns<sup>53</sup>) reported that cancer-free survival was higher after RFA than after PEI at 1, 2 and 3 years [e.g. 3-year rate (tumours 2.1–3 cm): RFA 40% vs. low-dose PEI 30% vs. high-dose PEI 32%<sup>59</sup>].

Five RCTs reported recurrence<sup>44,48,53,55,71</sup> or local tumour progression.<sup>59</sup> The outcome measures reported differed between RCTs (e.g. distant intrahepatic recurrence,<sup>44</sup> local recurrence,<sup>48,71</sup> etc). In the five better-quality RCTs (low RoB or some concerns), recurrence or local tumour progression occurred in more patients in the groups that received PEI,<sup>44,53,55,59,71</sup> although the difference was only small in one RCT (distant intrahepatic recurrence: RFA 32/70 vs. PEI 35/69<sup>44</sup>). One of these RCTs reported results by tumour size. Local tumour progression was similar between groups for smaller tumours (1–2 cm diameter) (3-year rate: RFA 9% vs. low-dose PEI 13% vs. high-dose PEI 12%), but it occurred in more patients with larger tumours (2.1–3 cm) after PEI treatment (RFA 18% vs. low-dose PEI 37% vs. high-dose PEI 33%).<sup>59</sup> In one low-quality RCT, the rate of local recurrence was similar between the two arms (5-year rate: RFA 11.7% vs. PEI 12.8%).<sup>48</sup>

Four RCTs reported a higher proportion of patients achieving complete response or complete ablation with RFA treatment than with PEI treatment.<sup>44,53,65,71</sup>

Findings on AEs were mixed, with some RCTs reporting worse AEs after RFA<sup>53,55,59,71</sup> and others reporting similar rates between treatment groups.<sup>44,48</sup> One high-quality RCT<sup>44</sup> and one low-quality RCT<sup>48</sup> reported a similar rate of major complications in each arm (RFA 2/70 vs. PEI 2/69;<sup>44</sup> RFA 0.9% vs. PEI 1.9%<sup>48</sup>). The rate of treatment-emergent AEs was also similar in the high-quality RCT (RFA 14.3% vs. PEI 17.4%).<sup>44</sup> In two RCTs, serious AEs were uncommon but only occurred in the RFA group (1.9% vs. 0;<sup>59</sup> 4.8% vs. 0<sup>53</sup>). AEs were also worse in the RFA group in the other two RCTs (RFA 32 vs. PEI 19 events;<sup>71</sup> RFA 5.1% vs. PEI 2.6% grade ≥ III events<sup>55</sup>). One RCT reported only that there were no mortalities related to either treatment.<sup>65</sup>

Five RCTs reported economic outcomes.<sup>44,48,53,55,59</sup> Two RCTs reported the direct medical costs of the procedures (see [Appendix 4](#) for details).<sup>44,48</sup> Three RCTs reported average length of hospital stay, which was considerably longer for patients who received RFA in two RCTs (RFA 4.2 days vs. PEI 1.7 days;<sup>53</sup> RFA 4.4 days vs. low-dose PEI 1.6 days vs. high-dose PEI 2.1 days<sup>59</sup>), but considerably longer for patients who received PEI in one RCT (RFA 10.8 days vs. PEI 26.1 days<sup>55</sup>).

### Radiofrequency ablation versus percutaneous acid injection

Only one RCT compared RFA with PAI, and it was judged to have some RoB concerns ( $n = 187$  patients).<sup>53</sup> OS was better in the RFA arm than the PAI arm (3-year survival: 74% vs. 53%; 10/62 deaths vs. 15/63 deaths). Cancer-free survival (3-year rate: 43% vs. 23%) and recurrence (3-year rate: 14% vs. 31%; 8 vs. 17 local recurrence events) were also better after treatment with RFA. Complete response was achieved in a similar proportion of tumours in each group (RFA 96.1% vs. PAI 92.4%). However, three serious AEs occurred in the RFA group (4.8% of patients) and none in the PAI group. Mean length of hospital stay was longer for patients who received RFA than for those receiving PAI (4.2 days vs. 2.2 days).

### Radiofrequency ablation versus laser ablation

Three RCTs compared RFA with laser ablation, with all three assessed as having some RoB concerns ( $n = 132$  patients).<sup>57,61,77</sup> One RCT included patients with tumours  $\leq 4$  cm, but the mean tumour size was 2.4 cm.<sup>61</sup> One RCT included a subgroup of patients with tumours  $\leq 2.5$  cm.<sup>77</sup>

Only one of the RCTs reported survival or progression outcomes.<sup>61</sup> There were no deaths in either treatment group, but PFS (1-year rate: RFA 86% vs. laser ablation 54%) and local disease progression (2/15 patients vs. 6/15 patients) were better in the RFA group than the laser ablation group.

Two RCTs reported complete response or complete ablation. In one RCT the proportion of tumours with complete ablation was higher in the RFA arm after both one procedure (86.7% vs. 66.7% of nodules) and two procedures (93% vs. 87%).<sup>61</sup> In the other RCT the complete response rate was similar between arms (RFA 92.3% vs. laser ablation 88.6%).<sup>57</sup>

One RCT measured patient satisfaction, using a self-made satisfaction questionnaire that included intraoperative discomfort, postoperative therapy effects, adverse reactions and physical recovery.<sup>57</sup> There was greater satisfaction with the laser ablation treatment than with RFA [great satisfaction (score 61–100 out of 100): RFA 64.1% vs. laser ablation 85.7%]. 30.8% of patients were dissatisfied (score < 60) after RFA, compared with just 5.7% of patients who received laser ablation.

All three RCTs reported AE results. One reported considerably more AEs (intra- or post-procedural) in patients who received RFA (93.3% vs. 13.3%), although there were no major complications in either arm.<sup>61</sup> In one RCT, postoperative rates of fever, nausea, vomiting, diarrhoea, abdominal pain and skin rash were similar between the two treatments.<sup>57</sup> The other RCT reported no major or minor complications during the procedures in either group.<sup>77</sup>

### Radiofrequency ablation versus resection

Seven completed RCTs compared RFA with surgical resection ( $n = 912$  patients).<sup>46,50,58,67,69,75,79</sup> One ongoing RCT was also identified.<sup>80</sup> One RCT did not report any data for the relevant subgroup (HCC  $\leq 2$  cm; the full population included patients with tumours < 4 cm, and the proportion with tumours  $\leq 3$  cm was not stated) and so RoB was not assessed.<sup>58</sup> Another RCT, which was judged to have some RoB concerns, did not report any relevant data for the  $\leq 3$  cm subgroup other than a KM curve.<sup>67</sup> Of the remaining RCTs, two were judged to have low RoB,<sup>75,79</sup> two had high RoB<sup>50,69</sup> and one had some concerns.<sup>46</sup> One of the RCTs recruited patients with recurrent HCC.<sup>75</sup>

Four RCTs reported OS,<sup>46,69,75,79</sup> with mixed findings. In one high-quality RCT<sup>79</sup> and one low-quality RCT,<sup>69</sup> OS at 1, 3 and 5 years was better after surgical resection [5-year rate: RFA 69% vs. resection 76%;<sup>79</sup>

5-year rate: RFA 61.4%/45.2% (solitary tumours/multifocal tumours) vs. resection 82.2%/69.2%<sup>69</sup>]. However, the RCT with some RoB concerns reported slightly better OS at 1, 2 and 3 years in the group that received RFA (3-year rate: 82.5% vs. 77.5%).<sup>46</sup> The high-quality RCT of recurrent HCC found that the two treatments were similar [HR (RFA vs. resection) 1.05, 95% CI 0.67 to 1.65].<sup>75</sup>

There were also mixed findings from the four RCTs that measured disease- or recurrence-free survival.<sup>46,50,75,79</sup> Recurrence-free survival was similar between treatment groups in one low-quality RCT (3-year rate: RFA 47.7% vs. surgery 49.8%; HR 0.96)<sup>50</sup> and the high-quality RCT of recurrent HCC patients (repeat-recurrence-free survival: HR 1.07, 95% CI 0.71 to 1.6<sup>75</sup>). The other high-quality RCT reported better disease-free survival after resection (5-year rate: 46% vs. 52%).<sup>79</sup> However, the disease-free survival rate was higher for patients who received RFA in the RCT with some RoB concerns (3-year rate: 55.4% vs. 41.3%).<sup>46</sup>

Only the RCT with some RoB concerns reported recurrence of HCC, with a similar proportion of patients experiencing recurrence in the RFA group as in the hepatectomy group (22/60 vs. 21/60).<sup>46</sup> This was also the only RCT to report on response, with a similar rate of complete tumour treatment after RFA as after surgery (57/60 vs. 58/60).

There were limited data on AEs reported. One RCT reported that postoperative complications (RFA 2/60 vs. resection 17/60), major complications (1/60 vs. 14/60) and serious pain requiring analgesia (3/60 vs. 43/60) were all more common after surgery than after RFA.<sup>46</sup> Four RCTs reported that there was no mortality related to the treatment or within the hospital admission period in either arm.<sup>46,50,58,69</sup>

Two RCTs reported average length of hospital stay, which was shorter for patients receiving RFA than resection in both RCTs (4 days vs. 7 days;<sup>79</sup> 4.3 days vs 11.8 days<sup>46</sup>). Length of intensive care unit (ICU) stay was also shorter after RFA (0 days vs. 6 days).<sup>46</sup>

### Radiofrequency ablation versus proton beam radiotherapy

One RCT of patients with recurrent or residual tumours compared RFA with proton beam radiotherapy and was judged to have some RoB concerns ( $n = 144$  patients).<sup>51</sup> OS was similar between the treatment groups (4-year rate: RFA 77.0% vs. proton beam radiotherapy 75.4%; HR at 2 years 1.07, 95% CI 0.58 to 1.98). PFS was also similar between treatment groups, with a median of 13.4 months after proton beam radiotherapy and 13.7 months after RFA. The rate of PFS was the same at 2 years (31.9% vs. 31.9%; HR 0.99, 95% CI 0.70 to 1.41), slightly higher after proton beam therapy at 3 years (17.9% vs. 26.3%), with a smaller difference between groups at 4 years (12.6% vs. 18.7%). The total number of progression events was greater in the RFA group (62/72 vs. 56/72). There were nine (16%) AEs at grade III or above in the RFA group compared with none in the proton beam radiotherapy group.

### Radiofrequency ablation versus radiofrequency ablation + transarterial chemoembolisation

Three RCTs compared RFA alone versus RFA combined with TACE ( $n = 220$  patients).<sup>43,54,74</sup> One included patients with recurrent HCC and was judged to be at low RoB.<sup>74</sup> The other two RCTs were at high RoB,<sup>43,54</sup> although one was only reported as a conference abstract, with very limited reporting of methods.<sup>43</sup>

All three RCTs reported OS. In the high-quality RCT of patients with recurrent HCC, OS was better at 1 and 3 years after RFA combined with TACE (3-year rate: RFA 60% vs. RFA + TACE 70%), but was the same in both arms at 5 years (50% vs. 50%).<sup>74</sup> Similarly, in one low-quality RCT, survival was better in the combined treatment arm at 2 years (88.8% vs. 100%) but similar by 3 and 4 years (4-year rate: RFA 74% vs. RFA + TACE 72.7%).<sup>54</sup> Overall the total number of deaths was similar between treatment arms in this RCT (5/46 vs. 6/43). However, in the other low-quality RCT, OS was better after treatment with RFA combined with TACE at 2 and 3 years (3-year rate: 73.9% vs. 84%), but similar at 1 year (RFA: 100% vs. RFA + TACE: 95.2%).<sup>43</sup>

Recurrence-free survival was higher after RFA combined with TACE in the high-quality RCT of patients with recurrent HCC (5-year rate: 26% vs. 48%).<sup>74</sup> One low-quality RCT reported higher local PFS in the combined treatment group at 2 years (74.1% vs. 81.1%) but a higher rate in the RFA-alone group at 3 and 4 years (4-year rate: 61.7% vs. 55.8%).<sup>54</sup> However, event-free survival (time from the beginning of treatment to last follow-up CT examination, local tumour progression, new lesions in the liver, distant metastasis, or death) was better after the combined treatment at 2, 3 and 4 years (4-year rate: 29.7% vs. 36.6%).

The two low-quality RCTs<sup>43,54</sup> both reported a similar rate of local tumour progression in both treatment groups (3-year rate: 8.7% vs. 9.5%;<sup>43</sup> 3-year rate: 14.4% vs. 17.6%<sup>54</sup>). Only one RCT reported response, with 100% of patients achieving complete response in both arms.<sup>54</sup> The rate of major complications was the same between treatment groups in the two low-quality RCTs.<sup>43,54</sup>

### **Radiofrequency ablation versus radiofrequency ablation + percutaneous ethanol injection**

Three RCTs compared RFA treatment alone versus RFA combined with PEI ( $n \geq 147$  patients).<sup>65,66,76</sup> One was judged to have a low RoB,<sup>76</sup> one some concerns<sup>65</sup> and one a high RoB.<sup>66</sup>

Overall survival was reported by two RCTs (one high quality and one low quality).<sup>66,76</sup> Both reported higher OS after treatment with RFA combined with PEI than after RFA alone (5-year rate: RFA 50.2% vs. RFA + PEI 55.3%;<sup>76</sup> 2-year rate: RFA 64.9% vs. RFA + PEI 79.0%<sup>66</sup>). In the low-quality RCT there was also more HCC recurrence after RFA treatment alone (2-year rate: 34.1% vs. 20.9%).<sup>66</sup>

Two RCTs reported data on response.<sup>65,76</sup> In both RCTs the rate of complete ablation was higher after one treatment of RFA combined with PEI than after one session of RFA alone. After two sessions of treatment (if necessary), the rate was similar between groups in the high-quality RCT<sup>76</sup> but remained higher in the RFA + PEI group in the RCT with some concerns (87.5% vs. 100%<sup>65</sup>).

Very limited data on AEs were reported. The two lower-quality RCTs reported that there were no serious AEs<sup>66</sup> or mortalities related to treatment<sup>65</sup> in either arm.

### **Radiofrequency ablation versus radiofrequency ablation + iodine-131 metuximab**

One RCT with some RoB concerns compared RFA alone with RFA and iodine-131 metuximab but reported limited data for the relevant subgroup of patients with tumours < 3cm ( $n = 78$  patients).<sup>78</sup> There was less recurrence in the group that received RFA combined with iodine-131 metuximab (HR 0.46, 95% CI 0.21 to 1.01). There were no serious AEs or treatment-related deaths in either group.

### **Radiofrequency ablation versus radiofrequency ablation + iodine-125**

One RCT with a low RoB compared RFA alone versus RFA and iodine-125 ( $n = 136$  patients).<sup>45</sup> OS was better after the combined treatment than after RFA alone (RFA: mean 70.8 months vs. RFA + iodine-125: 95.8 months; 36/68 vs. 23/68 deaths; HR 0.502, 95% CI 0.313 to 0.806). There was also less recurrence in patients who received the combined treatment (39/68 patients vs. 27/68 patients; HR 0.508, 95% CI 0.317 to 0.815; mean time to recurrence 66.8 vs. 93 months). Complete ablation was achieved in more patients with one treatment of RFA + iodine-125 than with one treatment of RFA alone, although after two treatments all participants in both arms had achieved complete ablation. There were more AEs at grade III or above after RFA combined treatment than after RFA alone (11 vs. 15 events; patient numbers not reported), although there were no procedure-related mortalities and no iodine-125 seed migration from the liver to the heart or other organs.

### **Radiofrequency ablation versus radiofrequency ablation + chemotherapy**

One RCT with a high RoB compared RFA alone versus RFA combined with chemotherapy ( $n = 38$  patients).<sup>47</sup> Recurrence was higher in the RFA group than in the RFA + chemotherapy group at 1 year (50% vs. 27%). There were no serious AEs in either group.

### Radiofrequency ablation + transarterial chemoembolisation versus resection

One RCT compared RFA combined with TACE versus partial hepatectomy, and it was judged to have some RoB concerns ( $n = 135$  patients).<sup>72</sup> The paper did not report any relevant efficacy data for the subgroup of patients with tumours  $\leq 3$  cm. However, KM curves for OS and recurrence showed that hepatectomy was more effective than RFA + TACE. There was no 30- or 90-day mortality in either arm.

### Percutaneous ethanol injection versus percutaneous acid injection

One RCT compared PEI with PAI and was judged to have some RoB concerns ( $n = 187$  patients).<sup>53</sup> OS (3-year rate: PEI 51% vs. PAI 53%; number of deaths 17/62 vs. 15/63), cancer-free survival (3-year rate: 21% vs. 23%), recurrence (3-year rate 34% vs. 31%; number of events 19/55 vs. 17/58) and complete response (88.1% vs. 92.4%) were all similar between arms. No serious AEs were reported in either arm. The average length of hospital stay was also similar between PEI and PAI groups (1.7 days vs. 2.2 days).

### Percutaneous ethanol injection versus resection

One RCT with high RoB compared PEI with resection ( $n = 82$  patients).<sup>49</sup> There was a higher rate of OS in the PEI arm at 2 and 3 years (3-year rate 96.7% vs. 88.1%) but by 4 years it was similar (92.1% vs. 88.1%) and at 5 years it was higher in the resection arm (46.0% vs. 81.8%). PFS was higher after resection at 1, 2, 3 and 4 years (4-year rate: 44.6% vs. 56.2%) but was similar by 5 years (44.6% vs. 48.2%). There was more recurrence of HCC in the PEI group (18/40 vs. 15/42 patients). Three patients had adverse effects after PEI, but for the resection arm the paper only reported that there were no significant complications.

### Percutaneous ethanol injection versus radiofrequency ablation + percutaneous ethanol injection

One RCT compared PEI alone versus RFA combined with PEI and was judged to have some RoB concerns ( $n = 48$  tumours).<sup>65</sup> The only relevant data reported were on complete response. After both one and two treatment sessions, no tumours in the PEI arm had been completely ablated, compared with 93.8% and 100%, respectively, in the RFA + PEI arm. Only 81.25% of tumours in the PEI group achieved complete ablation after all sessions. There were no mortalities related to either treatment.

### Percutaneous ethanol injection versus percutaneous ethanol injection + transarterial chemoembolisation

Two RCTs compared PEI alone with PEI combined with TACE ( $n = 82$  patients). One had a low RoB<sup>52</sup> and one had some bias concerns.<sup>60</sup> The two RCTs differed in their results. The high-quality RCT reported higher OS rates in the PEI + TACE arm at 1, 2 and 3 years (3-year rate: PEI 65.9% vs. PEI + TACE 80.8%), although it was similar between groups at 5 years (37.7% vs. 40.4%).<sup>52</sup> Rates of local residual disease (5-year rate 39.3% vs. 19.3%) and new nodular recurrence (5-year rate 100% vs. 50.2%) were lower after the combined PEI and TACE treatment. However, the lower-quality RCT reported a longer mean OS (57.2 vs. 42.4 months) and fewer deaths (6/14 vs. 8/13) in the PEI-alone arm.<sup>60</sup> Recurrence was also higher in the combined treatment arm (71.4% vs. 84.6%). However, the mean length of cancer-free survival was longer after PEI + TACE (16.7 vs. 22.9 months).<sup>60</sup>

The high-quality RCT reported two major complications (among 26 patients) in the combined treatment group and none in the PEI-alone group. Fever, continuous abdominal pain and transient increases in C-reactive protein were common AEs in both treatment groups.<sup>52</sup> The other RCT reported that no serious adverse effects or complications were related to either treatment.<sup>60</sup>

### Percutaneous acid injection versus percutaneous acid injection + transarterial chemoembolisation

One RCT with a high RoB compared PAI versus sequential TACE and PAI treatment ( $n = 55$  patients).<sup>70</sup> The rate of OS was 100% in both groups at 1 year, but at 3 years it was higher in the group that had received the combined treatment (49% vs. 73%). Data on cancer-free survival were not reported for the

subgroup of patients with tumours  $\leq 3$  cm (other than that there were no significant differences between treatment groups). There were no serious complications necessitating intensive care in either group.

### Percutaneous local ablative therapy versus resection

One RCT with high RoB compared percutaneous local ablative therapy (RFA, followed by RFA/PEI for any residual tumour, and TACE if residual tumour still remained) with partial hepatectomy ( $n = 79$  patients).<sup>68</sup> The paper did not report any relevant data for the subgroup of patients with tumours  $\leq 3$  cm, other than a KM curve. However, it reported that there were no significant differences in OS or disease-free survival between the two treatment groups for the  $\leq 3$  cm subgroup.

### Microwave ablation + sorafenib versus resection

One RCT with a high RoB compared treatment with MWA combined with sorafenib versus surgical resection ( $n = 120$  patients).<sup>56</sup> Rates of OS and tumour-free survival were similar between the two treatments at 1, 3 and 5 years, but mean OS was longer in the MWA + sorafenib group than the resection group (64.6 vs. 51.2 months). However, at 5 years there had been more recurrence of HCC in the MWA + sorafenib group (38.3% vs. 18.3%). Pain, fever, abdominal bleeding and infection were all experienced by considerably more patients in the resection arm than the MWA + sorafenib arm (pain: MWA + sorafenib 23.3% vs. resection 63.3%; fever: 25% vs. 48.3%; abdominal bleeding: 3.3% vs. 11%; infection: 1.7% vs. 30%).

### Ongoing trials

The electronic searches for RCTs undertaken on 3 February 2021 identified four potentially relevant ongoing RCTs: the published protocol by Zhu *et al.*<sup>80</sup> and three clinical trial register records, for which no further information was available. The searches for studies in progress and unpublished research, undertaken on 27 April 2021, identified 121 records in ClinicalTrials.Gov and 64 records in the European Union Clinical Trials Register; there was only one further potentially relevant ongoing RCT, after deduplication between databases. Further details are presented in [Table 5](#).

**TABLE 5** Table of potentially relevant ongoing RCTs

Study	Further details
<b>Ongoing studies identified from searches of bibliographic databases for RCTs</b>	
Zhu, 2021 <sup>80</sup>	Published protocol for a single centre (The Ninth People's Hospital of Chongqing, China) RCT comparing RFA vs. laparoscopic hepatectomy for small HCC (three or fewer tumours $\leq 3$ cm in diameter).
ClinicalTrials.gov: NCT04727307	Clinical trial register record describing a multicentre RCT comparing atezolizumab + bevacizumab combined with RFA vs. RFA alone for small HCC (one to three nodules $< 3$ cm). Sponsor: University Hospital, Montpellier, France. Actual study start date: 26 January 2021. Estimated primary completion date: January 2025. Estimated study completion date: July 2027.
ClinicalTrials.gov: NCT03790059	Clinical trial register record describing a multicentre RCT comparing RFA combined with recombinant human adenovirus Type 5 (H101) injection vs. RFA alone for small HCC (single lesion $\leq 3$ cm in diameter). Sponsor: Southwest Hospital, China. Study start date: October 2016. Estimated primary completion date: September 2020. Estimated study completion date: September 2020.
ClinicalTrials.gov: NCT04235660	Clinical trial register record describing a single-centre pilot RCT comparing Y90 radioembolisation vs. stereotactic body radiation therapy for solitary early-stage ( $\leq 3$ cm) HCC. Sponsor: Indiana University. Actual study start date: 22 July 2020. Estimated primary completion date: May 2024. Estimated study completion date: May 2024.
<b>Studies identified from searches of ClinicalTrials.Gov and the European Union Clinical Trials Register for ongoing RCTs</b>	
ClinicalTrials.gov: NCT04663035	Clinical trial register record describing a single-centre RCT comparing ablation followed by tislelizumab (immunotherapy) vs. ablation alone for early recurrent HCC. Sponsor: Sun Yat-sen University. Actual study start date: 21 December 2020. Estimated primary completion date: December 2023. Estimated study completion date: December 2025.

## Network meta-analysis results

Randomised controlled trials assessing the clinical effectiveness of ablative and non-surgical therapies for patients with early or very early HCC have been discussed and summarised in [Systematic review of RCTs](#). Four NMA models were produced, for the outcomes OS, PFS and overall recurrence and local recurrence.

Of the 37 RCTs described in [Systematic review of RCTs](#), six did not report any relevant data for the subgroup of patients with tumours  $\leq 3$  cm that could be included in the NMA<sup>57,58,62,64,65,78</sup> and one was ongoing, so no results were available.<sup>80</sup> A further three RCTs of patients with recurrent/residual HCC were not included.<sup>51,74,75</sup> Not all the resulting 27 RCTs included in the NMAs reported data for all four NMA outcomes; which RCTs reported for each outcome, as well as the type of data reported, are presented in [Report Supplementary Material 3](#).

Due to the small number of RCTs in each network, there was little evidence to inform the between-study heterogeneity. The uniform (0,3) prior distribution was considered in exploratory analyses and found to be too influential on the results. The half-normal (0, 0.19<sup>2</sup>) was used instead, as it expresses the prior belief that 95% of trials will give HRs within a factor of 2 from the estimated mean HR.<sup>28</sup> Results estimated using the half-normal (0, 0.50<sup>2</sup>) prior distribution are also reported.

Results for checks on the proportional hazards assumption are presented in [Report Supplementary Material 2](#). Schoenfeld residuals<sup>81</sup> were calculated for RCTs that reported the numbers at risk for the included KM curve. For RCTs that did not report the numbers at risk,<sup>46,77,82</sup> the proportional hazards assumption was assessed by visual inspection of the KM curves. For two trials (Aikata *et al.*<sup>43</sup> and Izumi *et al.*<sup>50</sup>) the proportional hazards assumption could not be tested as there were no KM curves available.

There were four RCTs for which the KM curves for OS crossed over,<sup>46,48,52,70</sup> which suggests that there may be some concerns about the proportional hazards assumption; however, for all other RCTs there was no statistical evidence that the assumption was violated. The validity of the NMAs depends on the proportional hazards assumption being correct, and more complex models with non-proportional hazards could not be fitted due to limitations of the data. Therefore, results should be interpreted with caution.

### Overall survival

#### Data

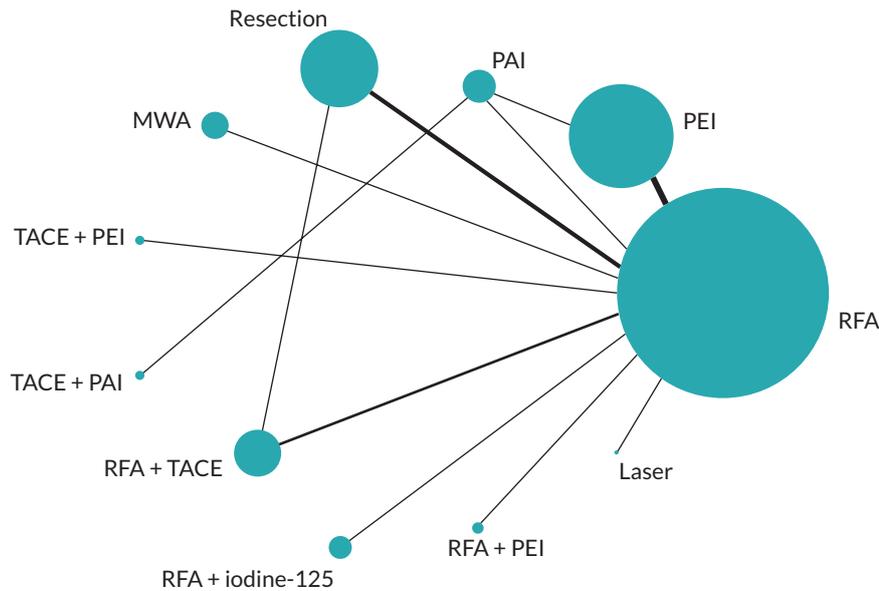
Of the 27 RCTs that reported relevant data, 16 were included in the NMA for OS. Eleven RCTs were excluded from the NMA: two<sup>47,50</sup> did not report OS data, and eight<sup>49,56,59,60,66-68,71</sup> reported data that would require strong assumptions to be made in order to calculate log-HRs required for the NMA; Orlicchio *et al.* (2014)<sup>61</sup> was also not included in the NMA as both arms in the trial reported zero deaths. Further details about the inclusion/exclusion of studies and how the evidence reported in the studies was transformed to a form suitable for NMA are summarised in the [Report Supplementary Material 3](#).

The network diagram for OS is presented in [Figure 2](#). Fifteen two-arm trials and one three-arm trial provided evidence on 11 interventions. A summary of the data used for the NMA is provided in [Report Supplementary Material 4](#).

#### Model selection and consistency checking

Model fit parameters for the FE and RE models are presented in [Report Supplementary Material 5](#). All three models fitted the data well, but as the difference in the DICs between the FE and RE models was  $< 3$ , the simpler FE model was chosen.

The 95% CrI for the RE model using the half-normal (0, 0.50<sup>2</sup>) prior was almost twice as wide as the 95% CrI for the model using half-normal (0, 0.19<sup>2</sup>), evidence that the priors for heterogeneity are



**FIGURE 2** Network diagram for OS.

Treatment nodes in the network diagram are scaled proportional to the number of patients who receive a particular treatment. The widths of the lines joining two nodes are weighted according to the number of studies that provide evidence on that comparison.

influential due to few studies being included for each comparison in the network. Plots for the prior and posterior distributions of the between-study heterogeneity for the RE models are presented in [Report Supplementary Material 5](#).

There was no evidence to suggest inconsistency in the network. Details of the inconsistency check and node-splitting results are presented in [Report Supplementary Material 5](#).

### Model results

Hazard ratios for OS for all treatments compared with RFA are presented in [Figure 3](#).

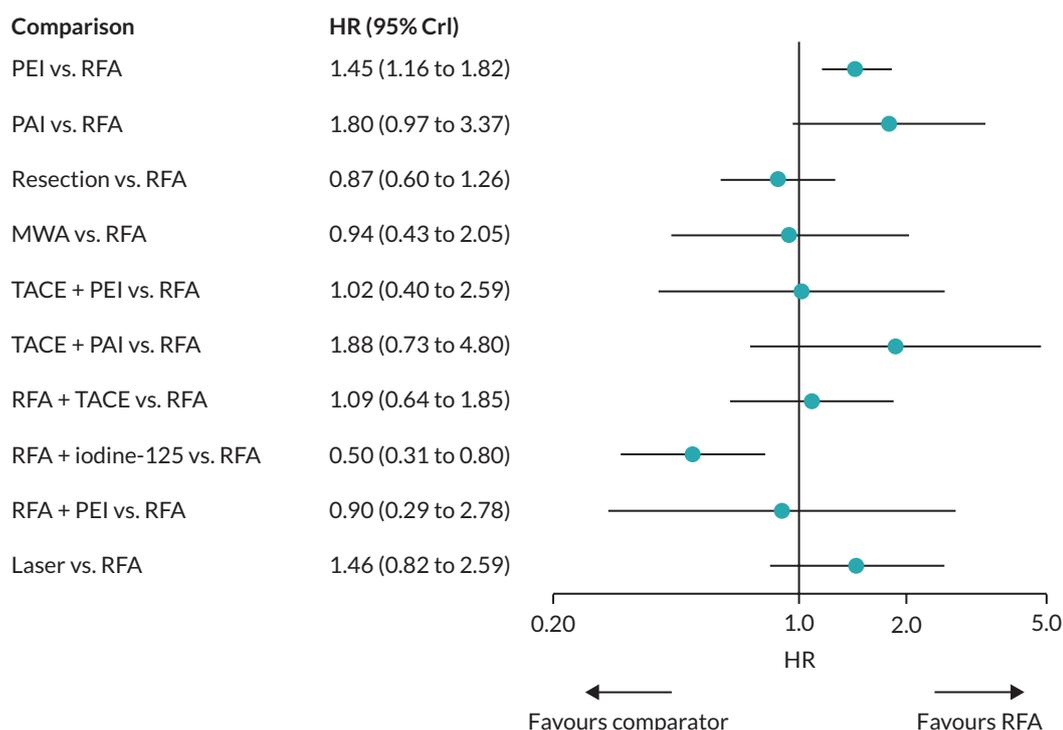
There was evidence to suggest that PEI worsens OS compared with RFA, and that RFA + iodine-125 improves OS compared with RFA (see [Figure 3](#)). There was also evidence to suggest that PEI worsens OS compared with resection, and that RFA + iodine-125 improves survival compared with PEI, PAI, TACE + PAI, RFA + TACE, and laser. There was insufficient evidence to suggest a difference in OS for all other treatment comparisons.

Hazard ratios comparing all treatment groups against each other for FE and RE models are reported in [Report Supplementary Material 5](#). Results for RE models displayed more uncertainty than the FE model, where results estimated using the wider half-normal (0, 0.50<sup>2</sup>) prior were more uncertain compared with results estimated using a half-normal (0, 0.19<sup>2</sup>) prior.

The mean and median ranks for each treatment, with their corresponding 95% CrIs, are presented in [Table 6](#). RFA + iodine-125 had the highest probability of being ranked the best treatment. However, there was a high level of uncertainty in treatment rankings; all treatments apart from RFA + iodine-125 displayed very wide CrIs. In fact, MWA, RFA + PEI, and TACE + PEI had 95% CrIs that included all 11 potential treatment ranks.

The treatment rank plot for OS (see [Figure 4](#)) also shows that RFA + iodine-125 had the highest probability of being the best treatment; however, the uncertainty in treatment ranks is also evident, as the probability of all other treatment ranks is < 50%.

## RESULTS



**FIGURE 3** Plot of HRs for OS compared with RFA for the FE model. HRs < 1 favour the comparator treatment over RFA.

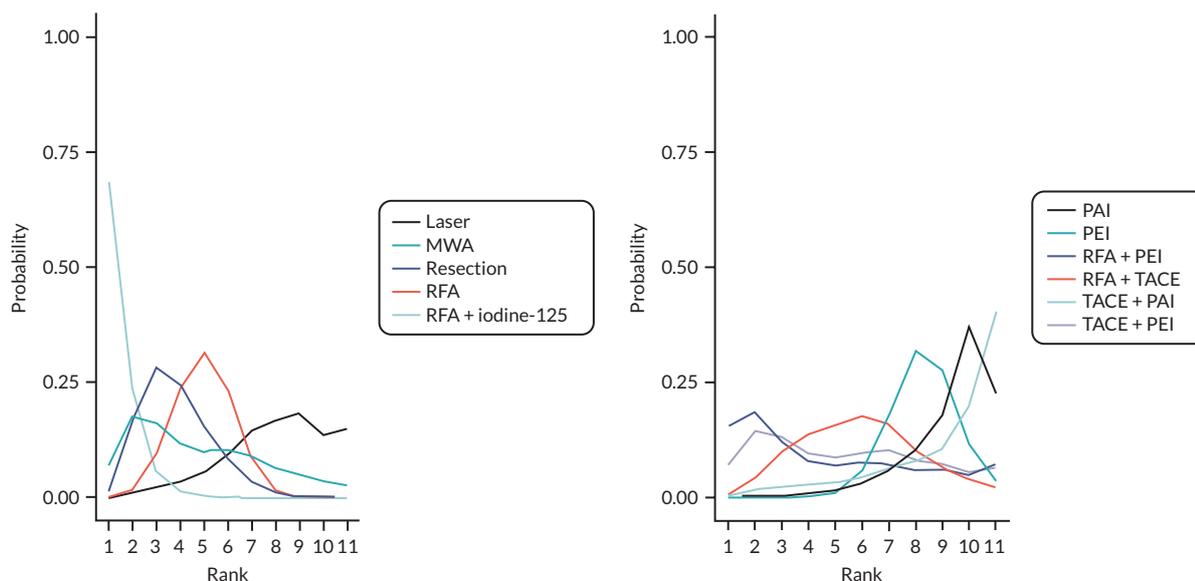
**TABLE 6** Mean and median treatment ranks for the FE model, with corresponding 95% CrIs for OS, sorted by mean rank out of 11 treatments

Treatments	Mean rank	Median rank	95% CrI for the rank
RFA + iodine-125	1.42	1	(1.00 to 3.00)
Resection	3.84	4	(2.00 to 7.00)
MWA	4.81	4	(1.00 to 11.00)
RFA + PEI	4.82	4	(1.00 to 11.00)
RFA	4.98	5	(3.00 to 7.00)
TACE + PEI	5.42	5	(1.00 to 11.00)
RFA + TACE	5.90	6	(2.00 to 10.00)
Laser	8.07	8	(3.00 to 11.00)
PEI	8.28	8	(6.00 to 11.00)
TACE + PAI	9.11	10	(3.00 to 11.00)
PAI	9.34	10	(5.00 to 11.00)

### Progression-free survival

#### Data

Of the 27 RCTs that reported relevant data, six were included in the NMA for PFS. Twenty-one RCTs were excluded from the NMA: 14<sup>43-45,47,48,52,55,66,67,69,70,72,76,77</sup> did not report PFS data, and five<sup>49,56,59,60,68</sup> reported data that would require strong assumptions to be made in order to calculate log-HRs required for the NMA; a further two<sup>61,63</sup> were excluded as they only reported local disease-free survival/PFS. Details about the inclusion/exclusion of studies and how the evidence reported in the studies was transformed into a form suitable for NMA are summarised in [Report Supplementary Material 3](#).



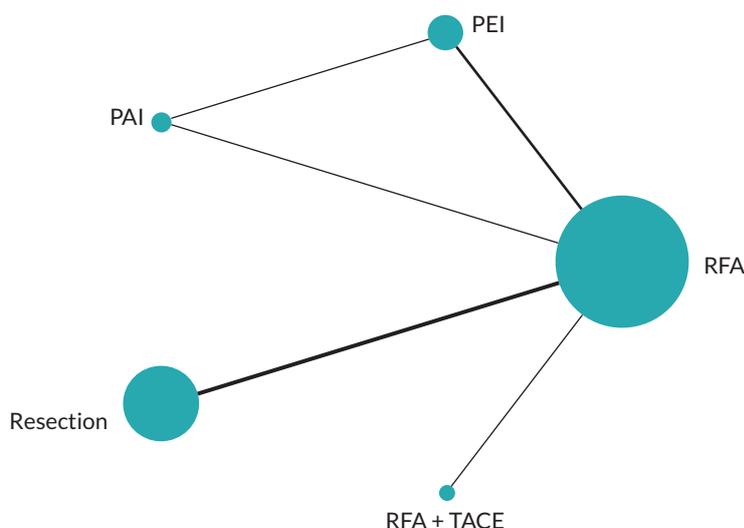
**FIGURE 4** Rank plot for OS for the FE model.

The network diagram for PFS is presented in [Figure 5](#). Five two-arm trials and one three-arm trial provided evidence on six interventions. A summary of the data used for the NMA is provided in [Report Supplementary Material 4](#).

### Model selection and consistency checking

Model fit parameters for the FE and RE models are presented in [Report Supplementary Material 5](#). All three models fit the data well, but as the difference in DICs between the fixed and RE models was  $< 3$ , the simpler FE model was chosen.

The between-study heterogeneity was low for the two RE models; however, the 95% CrI for the model using the half-normal  $(0, 0.50^2)$  prior was almost twice as wide as the 95% CrI for the model using the half-normal  $(0, 0.19^2)$  prior, evidence that the priors for heterogeneity are influential due to few studies



**FIGURE 5** Network diagram for PFS.

Treatment nodes in the network diagram are scaled proportional to the number of patients who receive a particular treatment. The widths of the lines joining two nodes are weighted according to the number of studies that provide evidence on that comparison.

## RESULTS

being included in the network. Plots for the prior and posterior distributions of the between-study heterogeneity for the RE models are presented in [Report Supplementary Material 5](#).

There is no potential for inconsistency in this network as there is no independent, indirect evidence for any of the comparisons – the single loop is formed by a three-arm study.<sup>53</sup>

### Model results

Hazard ratios for PFS for all treatments compared with RFA are presented in [Figure 6](#).

There was evidence to suggest that PEI and PAI are associated with worse PFS compared with RFA (see [Figure 6](#)). There was insufficient evidence to suggest a difference in PFS for all other treatment comparisons.

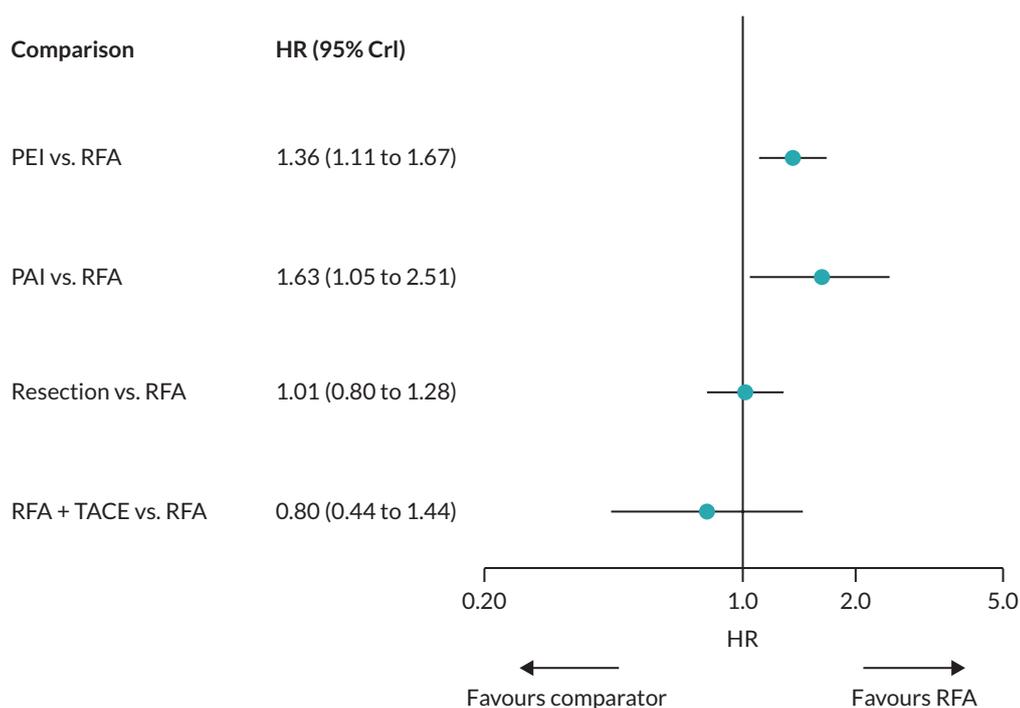
Hazard ratios comparing all treatment groups against each other for FE and RE models are reported in [Report Supplementary Material 5](#). Results for RE models displayed more uncertainty compared with the FE model, where results estimated using the wider half-normal (0, 0.50<sup>2</sup>) prior were more uncertain compared with results estimated using a half-normal (0, 0.19<sup>2</sup>) prior.

The treatment rank plot for PFS is presented in [Figure 7](#), and the mean and median ranks for each treatment, with their corresponding 95% CrIs are presented in [Table 7](#). RFA + TACE had the highest probability to be ranked the best treatment. However, there was a high level of uncertainty in the treatment ranking – all treatments displayed wide CrIs for ranks.

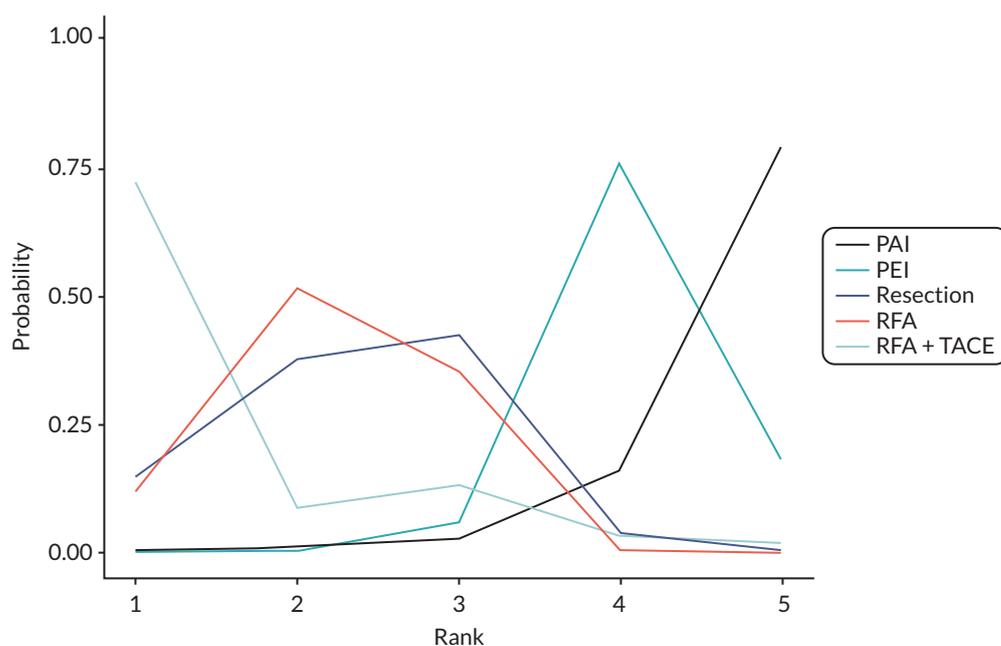
### Overall recurrence

#### Data

Of the 27 RCTs that reported relevant data, seven were included in the NMA for overall recurrence. Twenty RCTs were excluded from the NMA: 19<sup>43,48,50,52-54,59,61,63,66-72,76,77,79</sup> did not report overall recurrence data, and one reported distant recurrence.<sup>44</sup> Details about the inclusion/exclusion of RCTs and how the evidence reported was transformed into a form suitable for NMA are summarised in [Report Supplementary Material 3](#).



**FIGURE 6** Plot of HRs for PFS compared with RFA for the FE model. HRs < 1 favour the comparator treatment over RFA.



**FIGURE 7** Rank plot for PFS for the FE model

**TABLE 7** Mean and median ranks for the FE model, with the corresponding 95% CrIs for PFS, sorted by mean rank out of five treatments

Treatments	Mean rank	Median rank	95% CrI for the rank
RFA + TACE	1.53	1	(1.00 to 4.00)
RFA	2.24	2	(1.00 to 3.00)
Resection	2.38	2	(1.00 to 4.00)
PEI	4.12	4	(3.00 to 5.00)
PAI	4.73	5	(3.00 to 5.00)

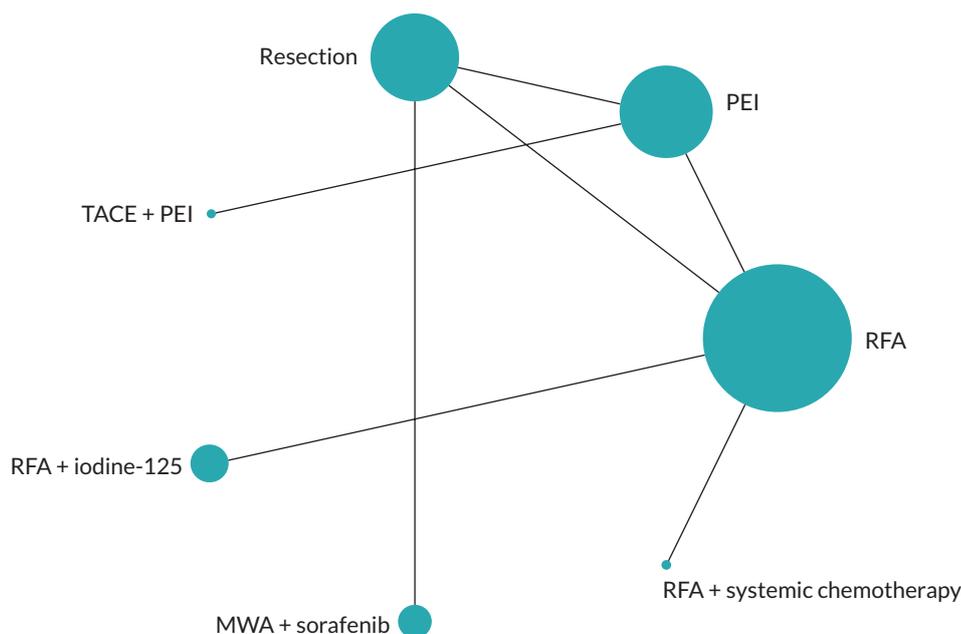
The network diagram for overall recurrence is presented in [Figure 8](#). Seven two-arm RCTs provided evidence on seven interventions. A summary of the data used for the NMA is provided in [Report Supplementary Material 4](#).

### Model selection and consistency checking

Model fit parameters for the FE and RE models are presented in [Report Supplementary Material 5](#). All three models fit the data well, but as the difference in DICs between the fixed and RE models was  $< 3$ , the simpler FE model was chosen.

The between-study heterogeneity was low for the two RE models; however, the 95% CrI for the model using the half-normal (0, 0.50<sup>2</sup>) prior was almost twice as wide as the 95% CrI for the model using the half-normal (0, 0.19<sup>2</sup>) prior, evidence that the priors for heterogeneity are influential due to few studies being included in the network. Plots for the prior and posterior distributions of the between-study heterogeneity for the RE models are presented in [Report Supplementary Material 5](#).

There was no evidence to suggest inconsistency in the network. Details of the inconsistency check and node-splitting results are presented in [Report Supplementary Material 5](#).



**FIGURE 8** Network diagram for overall recurrence.

Treatment nodes in the network diagram are scaled proportional to the number of patients who receive a particular treatment. The widths of the lines joining two nodes are weighted according to the number of studies that provide evidence on that comparison.

### Model results

Relative risks for overall recurrence for all treatments compared with RFA are presented in [Figure 9](#).

There was evidence to suggest that PEI increases the risk of overall recurrence compared with RFA (see [Figure 9](#)), and that RFA + iodine-125 decreases the risk of overall recurrence compared with RFA. The 95% CrIs of these estimates are very close to the 'null' effect.

There was evidence to suggest that RFA + iodine-125 decreases the risk of overall recurrence compared with PEI and TACE + PEI.

There was evidence to suggest that MWA + sorafenib increases the risk of overall recurrence compared with resection, and that RFA + iodine-125 and RFA + systemic chemotherapy decrease the risk of overall recurrence compared with MWA + sorafenib. There was insufficient evidence to suggest a difference in overall recurrence for all other treatment comparisons.

Relative risks comparing all treatment groups against each other for FE and RE models are reported in [Report Supplementary Material 5](#). Alternative models using arm-level data gave similar results. Results for RE models displayed more uncertainty compared with the FE model, where results estimated using the wider half-normal (0, 0.50<sup>2</sup>) prior were more uncertain compared with results estimated using a half-normal (0, 0.19<sup>2</sup>) prior.

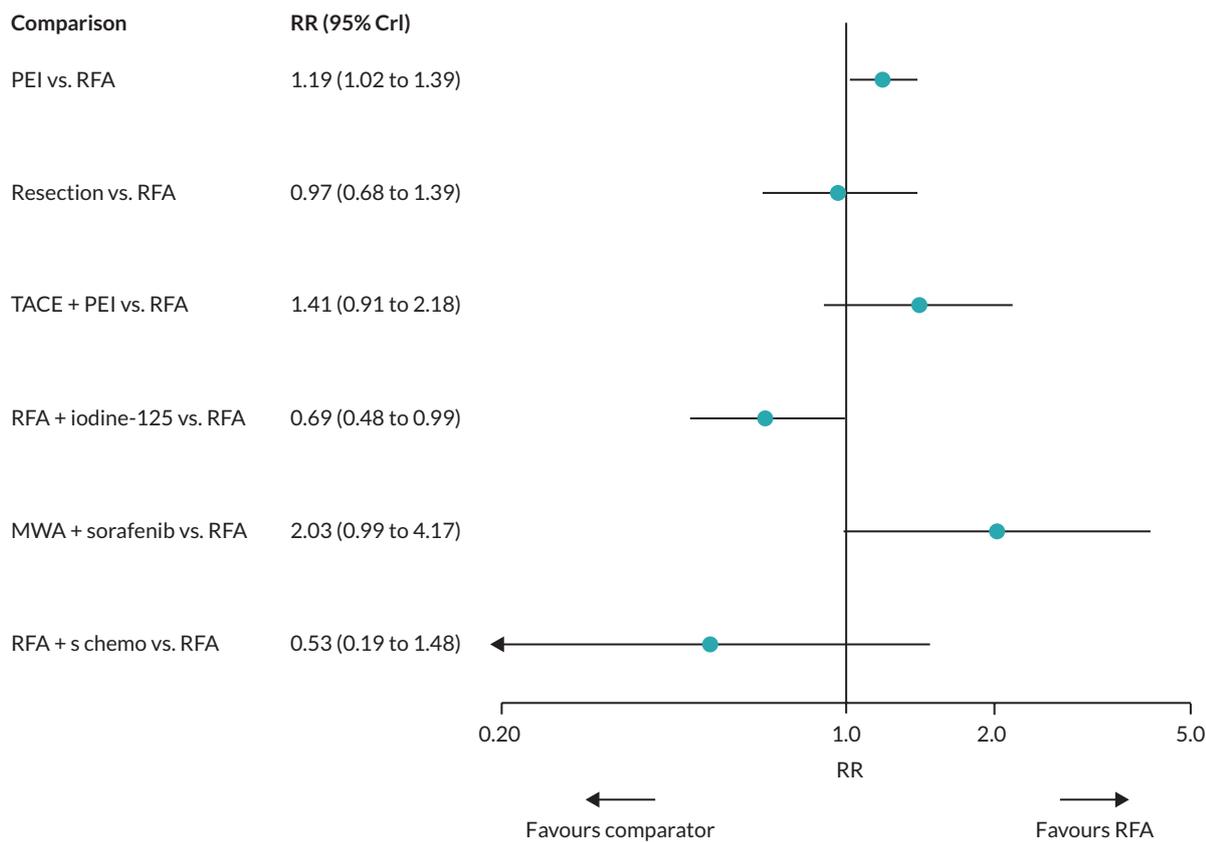
The treatment rank plot for overall recurrence is presented in [Figure 10](#), and the mean and median ranks for each treatment, with their corresponding 95% CrIs, are presented in [Table 8](#). RFA + systemic chemotherapy had the highest probability of being ranked the best. There was a high level of uncertainty in treatment rankings – all treatment ranks displayed wide CrIs.

### Local recurrence

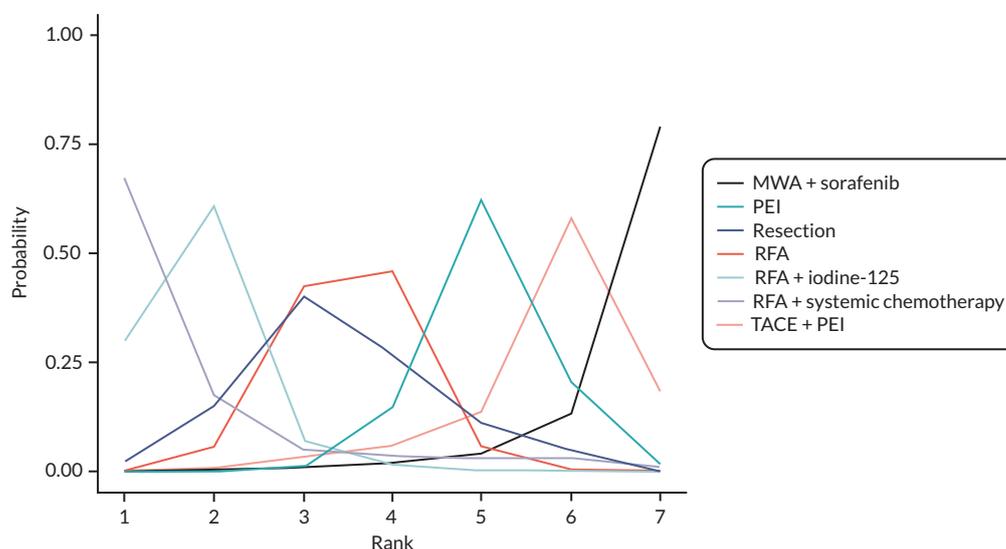
#### Data

Of the 27 RCTs that reported relevant data, 10 were included in the NMA for local recurrence.

Seventeen [44-47,49,50,55,56,60,67-70,72,76,77,79](#) did not report local recurrence data and were therefore excluded



**FIGURE 9** Plot of RRs for overall recurrence compared with RFA for the FE model. RRs < 1 favour the comparator treatment over RFA.



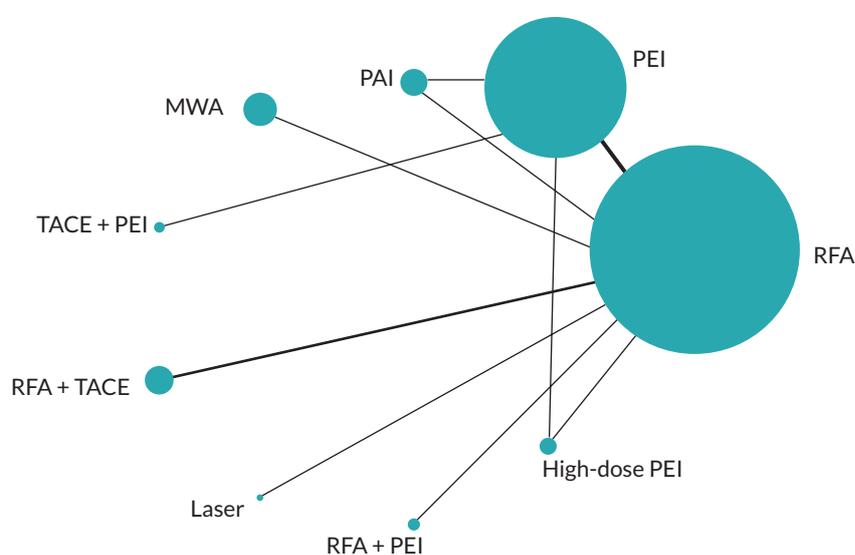
**FIGURE 10** Rank plot for overall recurrence for the FE model.

from the NMA. Details about the inclusion/exclusion of RCTs and how the evidence reported was transformed into a form suitable for NMA are summarised in [Report Supplementary Material 3](#).

The network diagram for overall recurrence is presented in [Figure 11](#). Eight two-arm and two three-arm RCTs provided evidence on nine interventions. A summary of the data used for the NMA is provided in [Report Supplementary Material 4](#).

**TABLE 8** Mean and median ranks for the FE model, with corresponding 95% CrIs for overall recurrence, sorted by mean rank, out of seven treatments

Treatments	Mean rank	Median rank	95% CrI for the rank
RFA + systemic chemotherapy	1.70	1	(1.00 to 6.00)
RFA + iodine-125	1.81	2	(1.00 to 3.00)
Resection	3.44	3	(2.00 to 6.00)
RFA	3.52	4	(2.00 to 5.00)
PEI	5.06	5	(4.00 to 6.00)
TACE + PEI	5.79	6	(3.00 to 7.00)
MWA + sorafenib	6.67	7	(4.00 to 7.00)

**FIGURE 11** Network diagram for local recurrence.

Treatment nodes in the network diagram are scaled proportional to the number of patients who receive a particular treatment. The widths of the lines joining two nodes are weighted according to the number of studies that provide evidence on that comparison.

### Model selection and inconsistency checking

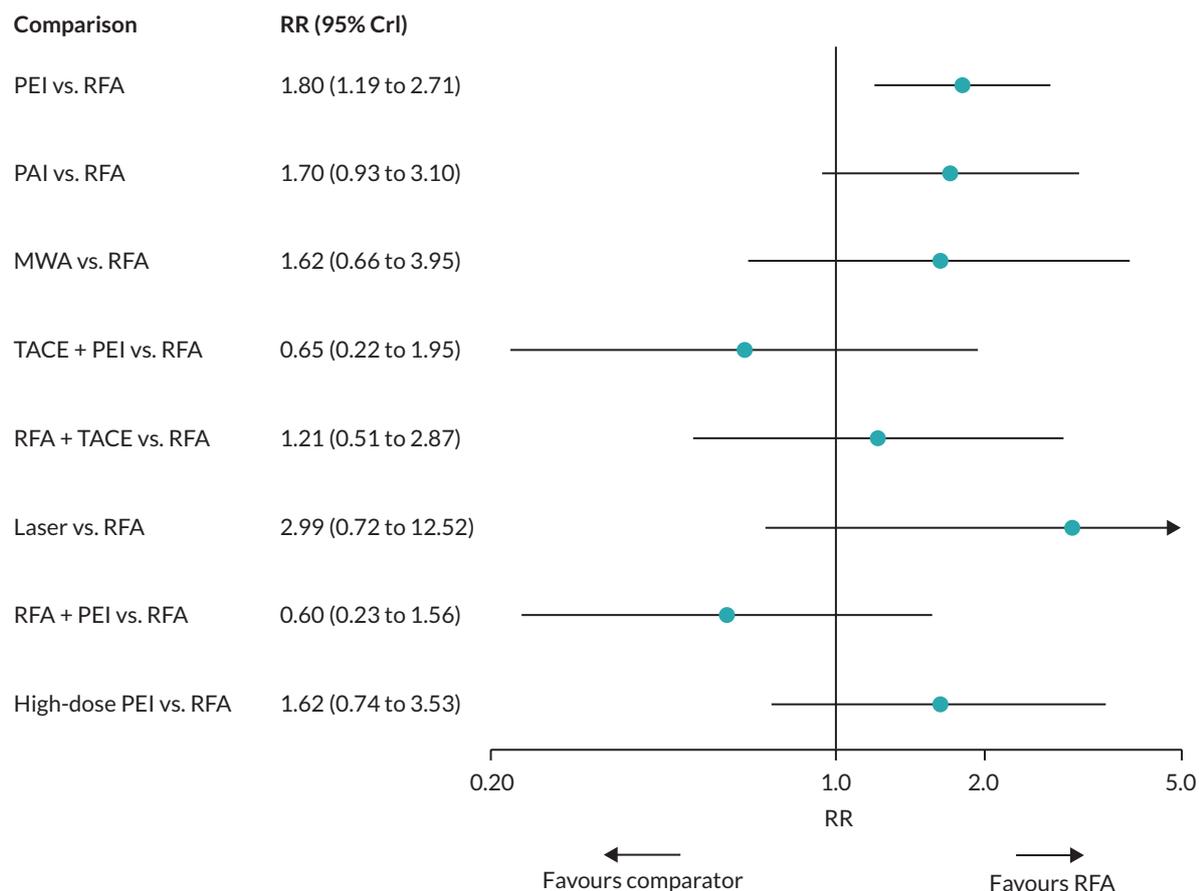
Model fit parameters for the FE and RE models are presented in [Report Supplementary Material 5](#). All three models fit the data well, but as the difference in DICs between the FE and RE models was  $< 3$ , a simpler FE model was chosen.

The between-study heterogeneity was low and consistent for the two RE models. However, the 95% CrI for the model using the half-normal  $(0, 0.50^2)$  prior was twice as wide as the 95% CrI for the model using the half-normal  $(0, 0.19^2)$ . Plots for the prior and posterior distribution of the between-study heterogeneity for the RE models are presented in [Report Supplementary Material 5](#).

There is no potential for inconsistency in this network as there is no independent, indirect evidence for any of the comparisons – the two loops in the network are formed by two separate three-arm studies.<sup>53,59</sup>

### Model results

Relative risks for local recurrence for all treatments compared with RFA are presented in [Figure 12](#).



**FIGURE 12** Plot of RRs for local recurrence compared with RFA for the FE model. RRs < 1 favour the comparator treatment over RFA.

There was evidence to suggest that PEI increases the risk of local recurrence compared with RFA (see [Figure 12](#)), and that RFA + PEI decreases the risk of local recurrence compared with PEI. There was insufficient evidence to suggest a difference in local recurrence for all other treatment comparisons.

Relative risks comparing all treatment groups against each other for FE and RE models are reported in [Report Supplementary Material 5](#). Alternative models using arm-level data gave similar results. Results for RE models displayed more uncertainty compared with the FE model, where results estimated using the wider half-normal (0, 0.50<sup>2</sup>) prior were more uncertain compared with results estimated using a half-normal (0, 0.19<sup>2</sup>) prior.

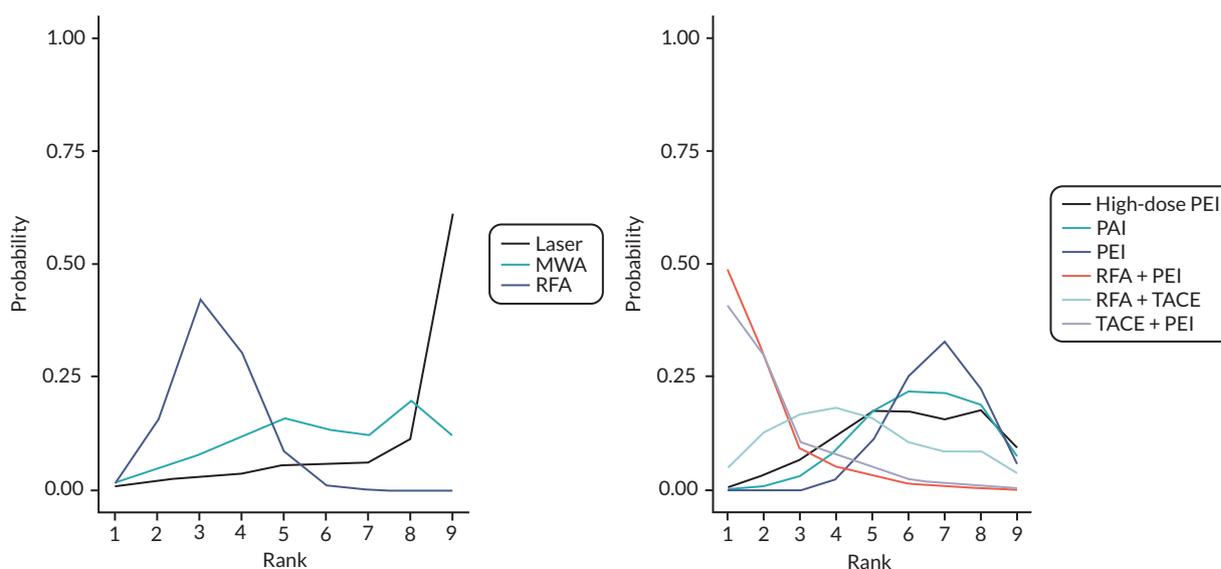
The treatment rank plot for local recurrence is presented in [Figure 13](#), and the mean and median ranks for each treatment, with their corresponding 95% CrIs, are presented in [Table 9](#). RFA + PEI had the highest probability of being ranked the best, although this probability was < 50%. The level of uncertainty in the treatment ranks was high – all treatments, with the exception of laser for the ninth rank, had rank probabilities below 50%. All treatments also had very wide CrIs for their rank.

## Threshold analysis of RCT networks

### Overall survival

The forest plot for the threshold analysis is presented in [Figure 14](#).

## RESULTS



**FIGURE 13** Rank plot for local recurrence for the FE model.

**TABLE 9** Mean and median ranks for the FE model, with corresponding 95% CrIs for local recurrence, sorted by mean rank, out of nine treatments

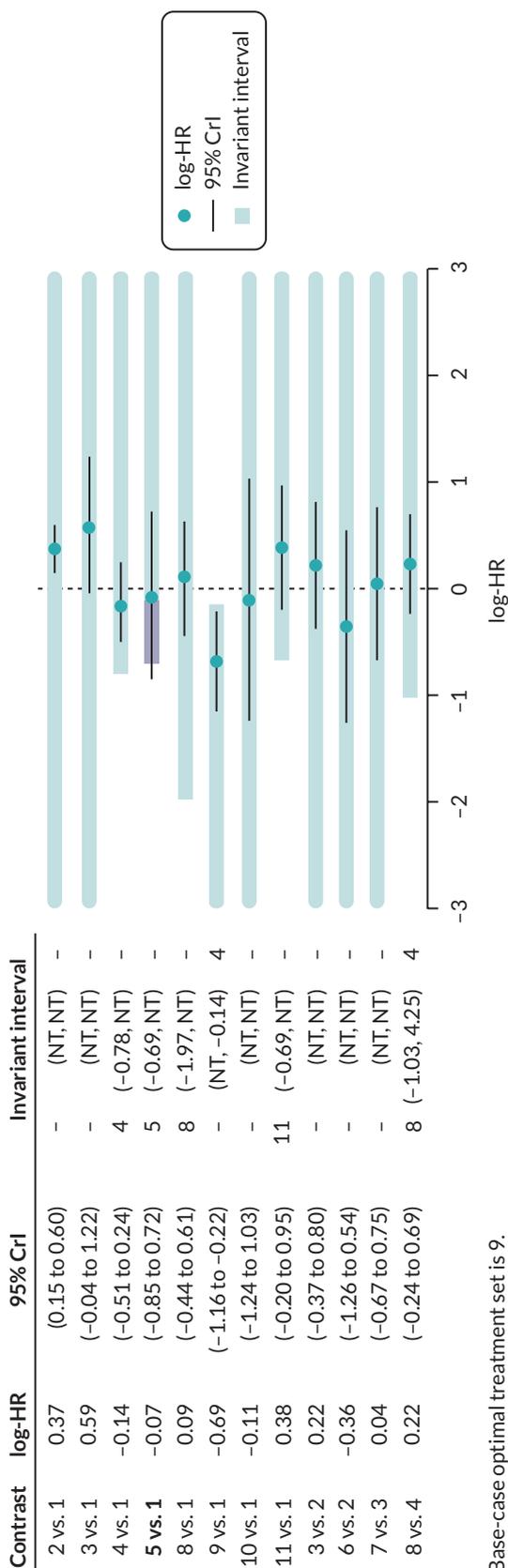
Treatments	Mean rank	Median rank	95% CrI of the rank
RFA + PEI	1.96	2	(1.00 to 6.00)
TACE + PEI	2.27	2	(1.00 to 7.00)
RFA	3.33	3	(2.00 to 5.00)
RFA + TACE	4.59	4	(1.00 to 9.00)
MWA	5.98	6	(2.00 to 9.00)
High-dose PEI	6.01	6	(2.00 to 9.00)
PAI	6.33	6	(3.00 to 9.00)
PEI	6.78	7	(4.00 to 9.00)
Laser ablation	7.75	9	(2.00 to 9.00)

Credible intervals for the MWA versus RFA (5 vs. 1) comparison extend beyond the limits of the invariance intervals, suggesting that the recommended treatment is sensitive to the uncertainty in the data.

As interventions that included PEI and PAI were not considered in the threshold analysis, comparisons including those interventions – PEI versus RFA (2 vs. 1), PAI versus RFA (3 vs. 1), RFA + PEI versus PEI (10 vs. 1), PAI versus PEI (3 vs. 2), TACE + PEI versus RFA (6 vs. 2), and TACE + PAI versus PAI (6 vs. 3) – had large thresholds on the log scale. None of the comparisons had thresholds that would be sensitive to small changes in log-HRs. The thresholds and new optimum treatments, based only on relative effects, are presented in [Report Supplementary Material 5](#).

### Progression-free survival

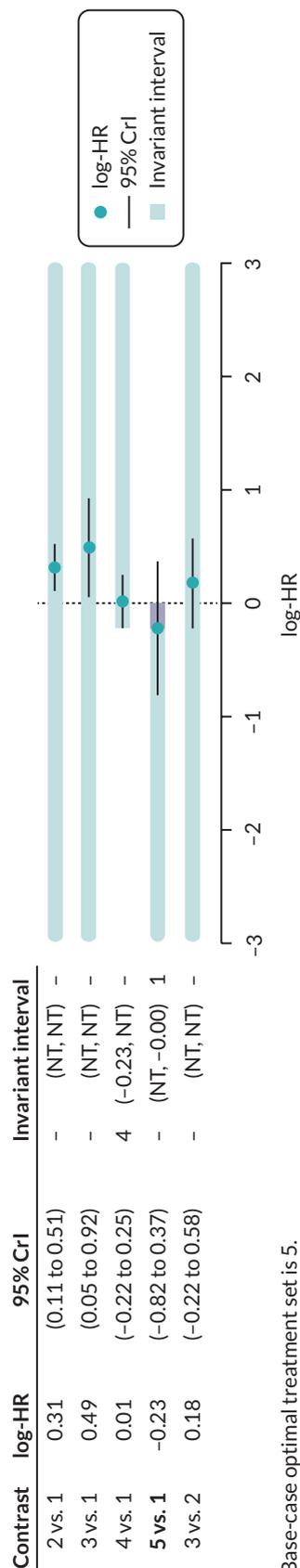
The forest plot for the threshold analysis is presented in [Figure 15](#).



**FIGURE 14** Forest plot for threshold analysis results for OS.

The light blue shaded bars represent the invariant interval; where the bars are shown in purple, CrIs for the comparison extend beyond the invariant interval.

**Treatment codes:** 1: RFA, 2: PEI, 3: PAI, 4: Resection, 5: MWA, 6: TACE + PEI, 7: TACE + PAI, 8: RFA + TACE, 9: RFA + iodine-125, 10: RFA + PEI, 11: Laser. The optimum treatment for this analysis was RFA + iodine-125.



**FIGURE 15** Forest plot for the threshold analysis for PFS.

The light blue shaded bars represent the invariant interval; where the bars are shown in purple, CrIs for the comparison extend beyond the invariant interval.

**Treatment codes:** 1: RFA, 2: PEI, 3: PAI, 4: Resection, 5: RFA + TACE. The optimum treatment for this analysis was RFA + TACE.

## RESULTS

Credible intervals for the RFA + TACE versus RFA (5 vs. 1) comparison extend beyond the limits of the invariance intervals, suggesting that the recommended treatment is sensitive to the uncertainty in the data, changing the optimum treatment to RFA.

Comparisons including PEI and PAI – PEI versus RFA (2 vs. 1), PAI versus RFA (3 vs. 1), and PAI versus PEI (3 vs. 2) – had very large thresholds on the log scale. However, the negative threshold for the resection versus RFA (4 vs. 1) comparison was very small, and a change of 0.24 units on the log-HR scale in the negative direction changes the optimum treatment to resection. Thresholds and new optimum treatments, based only on relative effects, are presented in [Report Supplementary Material 5](#).

### Overall recurrence

The forest plot for the threshold analysis is presented in [Figure 16](#).

Credible intervals for the RFA + iodine-125 versus RFA (5 vs. 1) and RFA + systemic chemotherapy versus RFA (7 vs. 1) comparisons extend beyond the limits of the invariance intervals, suggesting that the recommended treatment is sensitive to the uncertainty in the data.

Three comparisons – PEI versus RFA (2 vs. 1), resection versus PEI (3 vs. 2), and MWA + sorafenib versus resection (6 vs. 3) – had very large thresholds on the log scale. On the other hand, the negative threshold for the RFA + iodine-125 versus RFA (5 vs. 1) comparison was very small, and a change of 0.26 units on the log-RR scale in the negative direction changes the optimum treatment to RFA + iodine-125. Additionally, the positive threshold for RFA + systemic chemotherapy versus RFA (7 vs. 1) was very small, and a change of 0.26 units on the log-RR scale in the positive direction also changes the optimum treatment to RFA + iodine-125. Thresholds and new optimum treatments, based only on relative effects, are presented in [Report Supplementary Material 5](#).

### Local recurrence

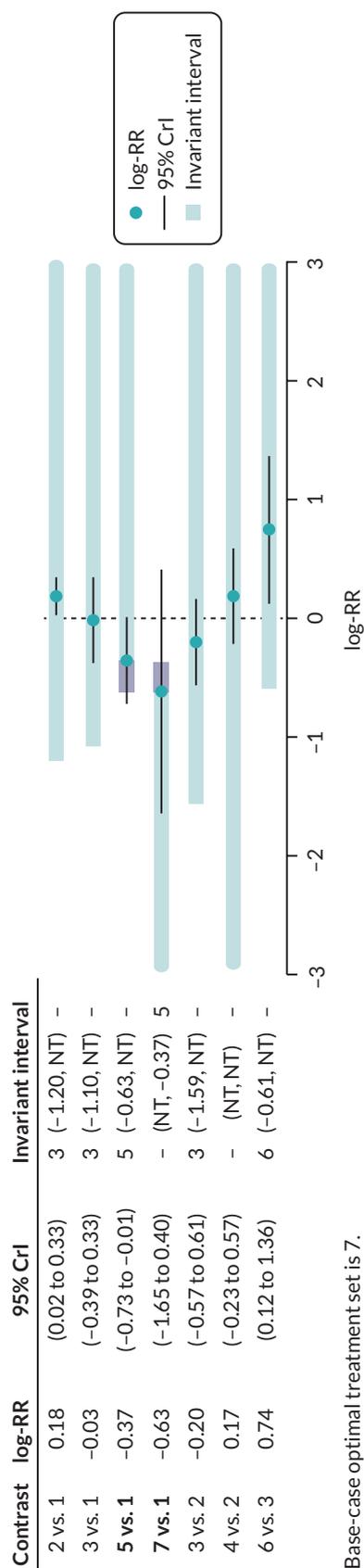
The forest plot for the threshold analysis is presented in [Figure 17](#).

Credible intervals for the MWA versus RFA (4 vs. 1), RFA + TACE versus RFA (6 vs. 1), and laser versus RFA (7 vs. 1) comparisons extend beyond the limits of the invariance intervals, suggesting that the recommended treatment is sensitive to the uncertainty in the data.

Comparisons including PEI and PAI – PEI versus RFA (2 vs. 1), PAI versus RFA (3 vs. 1), RFA + PEI versus RFA (8 vs. 1), high-dose PEI versus RFA (9 vs. 1), PAI versus PEI (3 vs. 2), TACE + PEI versus PEI (5 vs. 2), and high-dose PEI versus PEI (9 vs. 2) – had very large thresholds on the log-RR scale, as did the laser versus RFA (7 vs. 1) comparison. On the other hand, the negative threshold for the MWA versus RFA (4 vs. 1) and RFA + TACE versus RFA (6 vs. 1) comparisons was small, and changes of 0.48 and 0.19 units on the log-RR scale in the negative direction would change the optimum treatment to MWA and RFA + TACE, respectively. Thresholds and new optimum treatments, based only on relative effects, are presented in [Report Supplementary Material 5](#).

## Systematic review of non-randomised evidence

The electronic searches for non-randomised studies of selected interventions, where RCT evidence was not available (HIFU, cryoablation, IRE, ECT, histotripsy, SABR and wider radiotherapy techniques) or for comparisons where the threshold analysis suggested that additional evidence could plausibly change the NMA result (RFA, MWA and laser ablation, compared with each other or surgical resection), identified a total of 8009 records after deduplication between databases. One additional record was identified from screening reference lists of relevant systematic reviews. Clinical advisors were not aware of any additional studies, other than those already identified from the electronic searches. However, clinical advisors were aware of additional unpublished data from a prospective registry of patients undergoing

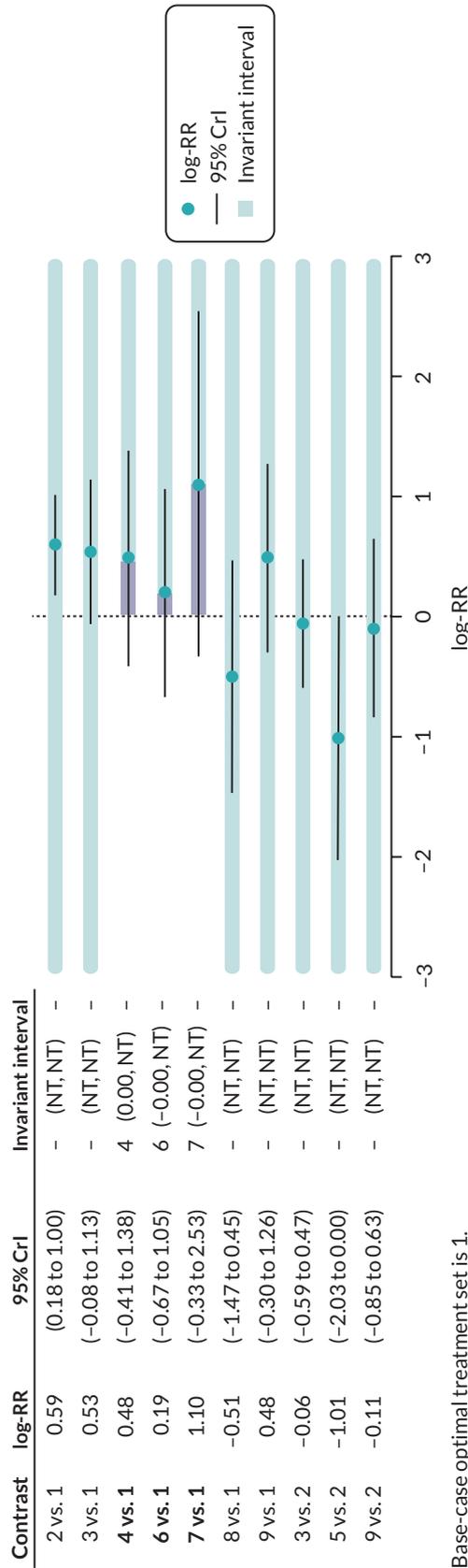


Base-case optimal treatment set is 7.

**FIGURE 16** Forest plot for threshold analysis results for overall recurrence.

The light blue shaded bars represent the invariant interval; where the bars are shown in purple, CrIs for the comparison extend beyond the invariant interval.

**Treatment codes:** 1: RFA, 2: PEI, 3: Resection, 4: TACE + PEI, 5: RFA + iodine-125, 6: MWA + sorafenib, 7: RFA + systemic chemotherapy. The optimum treatment for this analysis was RFA + systemic chemotherapy.



Base-case optimal treatment set is 1.

**FIGURE 17** Forest plot for results of threshold analysis for local recurrence.

The light blue shaded bars represent the invariant interval; where the bars are shown in purple, CrIs for the comparison extend beyond the invariant interval.

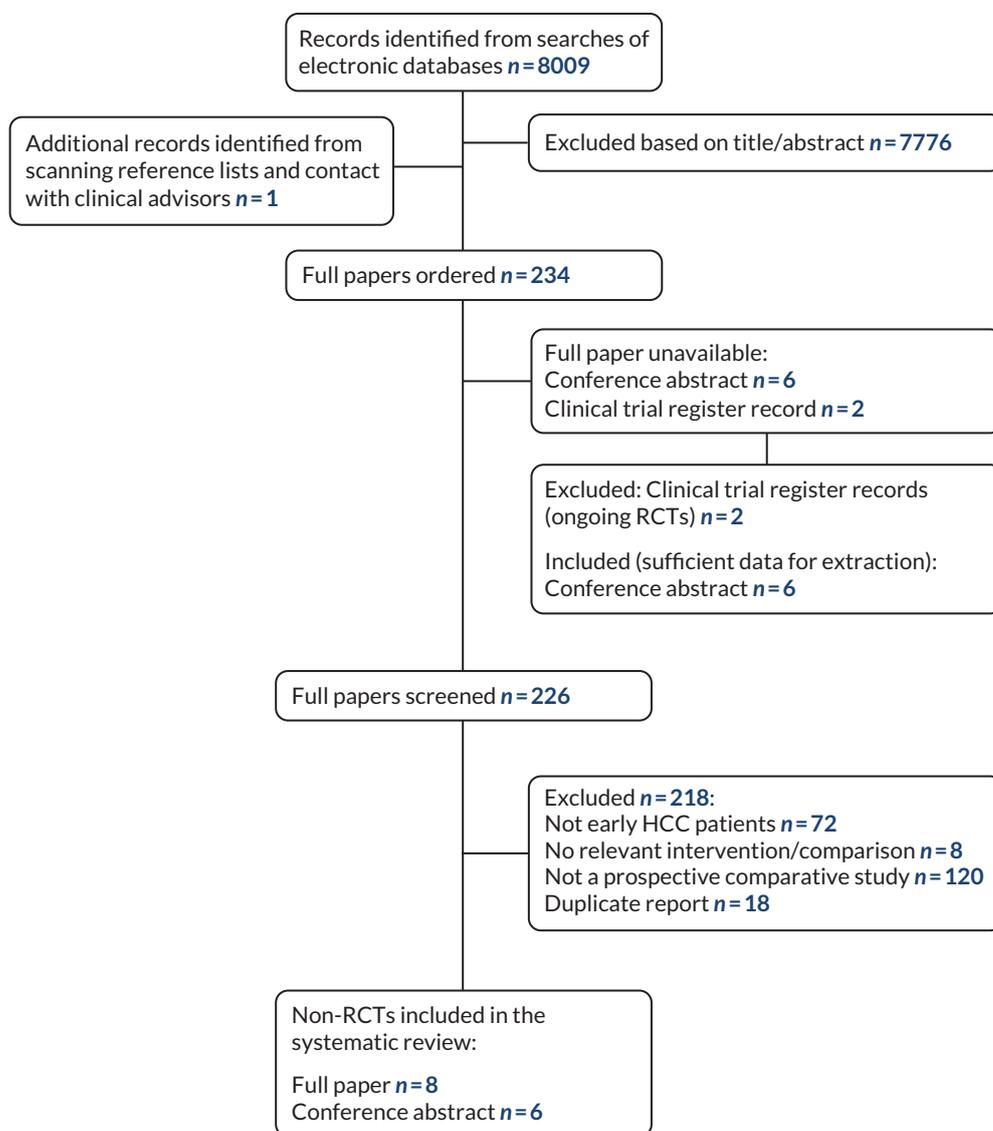
**Treatment codes:** 1: RFA, 2: PEI, 3: PAI, 4: MWA, 5: TACE + PEI, 6: RFA + TACE, 7: Laser, 8: RFA + PEI, 9: High-dose PEI. The optimum treatment for this analysis was RFA.

treatment for HCC at Leeds Teaching Hospitals NHS Trust. These data were made available for use in the updated NMAs (see [Updated network meta-analyses using RCT and non-RCT evidence](#)).

Two hundred and thirty-four potentially relevant studies were ordered for full paper screening. Eight papers were unavailable as they were only reported as conference abstracts or clinical trial register records. Two hundred and twenty-six full papers were screened; 218 were excluded at the full paper stage and are listed in [Appendix 5](#), along with the reasons for their exclusion. [Figure 18](#) presents the flow of non-RCT studies through the study selection process.

### Characteristics of non-randomised studies included in the review

Details of the 14 non-randomised comparative studies that were included in the systematic review are presented in [Table 10](#). Eight of the 14 studies restricted inclusion criteria to HCC patients with tumour size up to 3 cm in diameter.<sup>83-90</sup> One study restricted inclusion criteria to HCC patients with tumour size up to 2 cm in diameter.<sup>91</sup> One study included patients with tumours up to 5 cm in diameter, but reported separate results for the subgroup of patients with tumours up to 3 cm in diameter.<sup>92</sup> Two studies did not report specific tumour size inclusion criteria, but in one study average tumour size was 2.15 ( $\pm 0.53$ )



**FIGURE 18** Flow diagram of the study selection process (non-RCTs).

**TABLE 10** Non-RCTs included in the systematic review

Study	Location	Participant information	Intervention	Comparator
Barabino, 2016 <sup>93</sup> (abstract)	Italy	154 patients with HCC unsuitable for percutaneous treatments or hepatic resection (average tumour size 2.15 ( $\pm$ 0.53) cm in one arm and 1.92 ( $\pm$ 0.5) cm in the other)	Laparoscopic RFA	Laparoscopic MWA
Cheung, 2013 <sup>83</sup>	China	106 patients (with 119 tumours) with < 3 cm tumours (primary or first recurrence)	HIFU	RFA
Choi, 2004 <sup>84</sup> (abstract)	Korea	164 patients with $\leq$ 3 cm tumours	RFA	Hepatic resection
Du, 2012 <sup>92</sup> (reported in Chinese)	China	116 patients with tumours $\leq$ 5 cm; subgroup of 60 patients with tumours $\leq$ 3 cm	RFA	Surgical resection
Ei, 2015 <sup>95</sup>	Japan	119 patients with < 5 cm tumours, included a few patients with tumours > 3 cm; median 2.5 cm in cryoablation group (maximum 4 cm), median 1.9 cm in RFA/MWA group (maximum 4.5 cm)	Cryoablation	RFA or MWA
Elgendi, 2014 <sup>85</sup> (abstract)	Egypt	51 patients with < 3 cm tumours in locations not amenable for percutaneous route	Intraoperative RFA	Surgical resection
Elgendi, 2015 <sup>91</sup> (abstract)	Egypt	92 patients with < 2 cm tumours in locations not amenable for percutaneous route	Intraoperative RFA	Surgical resection
Harada, 2016 <sup>96</sup>	Japan	121 patients with < 5 cm tumours and portal hypertension, included a few patients with tumours > 3 cm in the resection group; mean 2.1 cm (range 0.7–5 cm)	RFA	Liver resection
Horigome, 2000 <sup>86</sup>	Japan	105 patients with $\leq$ 3 cm tumours	Resection	MWA PEI
Huang, 2014 <sup>87</sup>	China	346 patients with $\leq$ 3 cm tumours	RFA	Surgical resection
Peng, 2010 <sup>88</sup> (abstract)	China	195 patients with $\leq$ 3 cm tumours	RFA (n = 79), surgical resection (n = 24)	Surgical resection (n = 75), RFA (n = 17)
Qian, 2012 <sup>89</sup>	China	42 patients with < 3 cm tumours	MWA	RFA
Sugimoto, 2019 <sup>94</sup>	Japan	21 patients (with 24 tumours; median tumour size 2.03 (SD 0.44) cm in one arm and 1.73 (SD 0.67) cm in the other)	IRE	RFA
Tateishi, 2020 <sup>90</sup> (abstract)	Japan	740 patients with $\leq$ 3 cm tumours	RFA	Surgery

cm in one arm and 1.92 ( $\pm$  0.5) cm in the other,<sup>93</sup> and the other study reported median tumour size of 2.03 (SD 0.44) cm in one arm and 1.73 (SD 0.67) cm in the other.<sup>94</sup> Two studies included patients with tumours up to 5 cm, but a clear majority of patients had tumours < 3 cm in diameter.<sup>95,96</sup> In three of the included studies the patients had tumours unsuitable for percutaneous treatment,<sup>85,91,93</sup> and one study included patients with primary or first recurrent HCC.<sup>83</sup> Study sample sizes ranged from 21 to 740 patients.

The majority of studies were conducted in Asian countries, which has implications for the generalisability of results to the UK population, as discussed in [Characteristics of RCTs included in the review](#). Studies were

**TABLE 11** Number of non-RCTs making each comparison

RFA	RFA				
MWA	2	MWA			
PEI*		1	PEI <sup>a</sup>		
Resection	8	1	1	Resection	
HIFU	1			HIFU	
Cryoablation					Cryoablation
RFA/MWA					1 RFA/MWA
IRE	1				

a Note that PEI was included as part of a three-arm trial comparing MWA, resection and PEI; we did not search for non-RCTs of PEI.

conducted in China ( $n = 5$ ), Japan ( $n = 5$ ), Egypt ( $n = 2$ ), Korea ( $n = 1$ ) and Italy ( $n = 1$ ). Six of the included studies were only reported as conference abstracts, and therefore limited data were available.<sup>84,85,88,90,91,93</sup> While the inclusion criteria stated that only prospective studies were eligible for inclusion, for six studies it was not possible to determine whether patients were recruited prospectively or retrospectively; these studies were included to ensure that no relevant data were missed.<sup>83–85,91,93,95</sup>

**Table 11** shows the comparisons made in the included studies, 13 of which assessed RFA. While RFA was usually delivered via the percutaneous route, three studies assessed laparoscopic<sup>93</sup> or intraoperative RFA<sup>85,91</sup> in patients with tumours unsuitable for percutaneous treatment. It should also be noted that several of the studies allocated patients to treatment groups depending on their tumour characteristics. Cheung *et al.* offered HIFU to patients with poor liver function or decompensated cirrhosis or tumours located at sites considered difficult for RFA;<sup>83</sup> Ei *et al.* allocated patients to cryoablation if tumours were in close vicinity to major veins or organs;<sup>95</sup> both studies by Elgendi *et al.* allocated patients depending on the location and depth of the tumour from the liver capsule;<sup>85,91</sup> Harada *et al.* allocated patients depending on Child–Pugh class, tumour location and indocyanine green retention tests;<sup>96</sup> and Sugimoto *et al.* allocated patients depending on operator preference, tumour size, geometry and location.<sup>94</sup> In the study by Peng *et al.*, patients were allocated to RFA or surgical resection as the first choice, but the actual treatment received depended on the tumour location.<sup>88</sup>

### Quality of non-randomised studies included in the review

Results of the quality assessment of the non-randomised comparative studies are presented in **Table 12**. Six of the included studies were only reported as conference abstracts, and therefore there are a few 'Unclear' responses to some of the quality assessment criteria owing to the limited reporting.

Generally, methods were poorly reported. Inclusion criteria were clearly defined in 8/14 studies. The intervention was clearly described and consistently delivered in 8/14 studies and the comparator was clearly described and consistently delivered in 7/14 studies. None of the studies reported whether outcome assessors were blinded to treatment group.

Allocation to treatment groups was adequately described and appropriate in only two studies, resulting in patients having similar baseline characteristics between groups.<sup>87,89</sup> As discussed in **Overall survival**, several of the studies allocated patients to treatment groups depending on their tumour characteristics. Because appropriateness of treatment allocation and similarity of treatment groups at baseline were two of the important quality assessment criteria, this resulted in the other 12 studies having a high overall RoB judgement. The study by Qian *et al.* was the only study to have a low overall RoB judgement<sup>89</sup>

**TABLE 12** Risk of bias assessment results (non-RCTs)

Trial	Inclusion criteria clearly defined <sup>a</sup>	Allocation to treatment groups adequately described/ appropriate <sup>a</sup>	Groups similar at baseline <sup>a</sup>	Clearly described and consistently delivered intervention	Clearly described and consistently delivered comparator	Outcome assessors blinded	Missing outcome data balanced across groups <sup>a</sup>	Free from suggestion of selective reporting	Overall judgement of ROB
Barabino, 2016 <sup>93</sup> (abstract)	No	No	No	No	No	Unclear	Unclear	Unclear	High
Cheung, 2013 <sup>83</sup>	Yes	No	No	Yes	Yes	Unclear	Unclear	Yes	High
Choi, 2004 <sup>84</sup> (abstract)	No	No	Unclear	No	No	Unclear	Unclear	Unclear	High
Du, 2012 <sup>92</sup> (Chinese)	Yes	No	Unclear	Yes	Yes	Unclear	Yes	Yes	High
Ei, 2015 <sup>95</sup>	Yes	No	No	Yes	No	Unclear	Yes	Yes	High
Elgendi, 2014 <sup>85</sup> (abstract)	No	No	Unclear	No	No	Unclear	Unclear	Unclear	High
Elgendi, 2015 <sup>91</sup> (abstract)	No	No	Unclear	No	No	Unclear	Unclear	Unclear	High
Harada, 2016 <sup>86</sup>	Yes	No	No	Yes	Yes	Unclear	Yes	Yes	High
Horigome, 2000 <sup>86</sup>	No	No	No	Yes	Yes	Unclear	Yes	Yes	High
Huang, 2014 <sup>87</sup>	Yes	Yes	Yes	Yes	Yes	Yes for HRQoL, unclear for survival/ AE	Unclear	Yes	Unclear
Peng, 2010 <sup>88</sup> (abstract)	No	No	Unclear	No	No	Unclear	Unclear	Unclear	High
Qian, 2012 <sup>89</sup>	Yes	Yes	Yes	Yes	Yes	Unclear	Yes	Yes	Low
Sugimoto, 2019 <sup>94</sup>	Yes	No	No	Yes	Yes	Unclear	Unclear	Yes	High
Tateishi, 2020 <sup>90</sup> (abstract)	Yes	No	No	No	No	Unclear	Unclear	Yes	High
<b>Total</b>	<b>Yes: 8 No: 6 Unclear: 0</b>	<b>Yes: 2 No: 12 Unclear: 0</b>	<b>Yes: 2 No: 7 Unclear: 5</b>	<b>Yes: 8 No: 6 Unclear: 0</b>	<b>Yes: 7 No: 7 Unclear: 0</b>	<b>Yes: 0 No: 0 Unclear: 14</b>	<b>Yes: 5 No: 0 Unclear: 9</b>	<b>Yes: 9 No: 0 Unclear: 5</b>	<b>High: 12 Low: 1 Unclear: 1</b>

<sup>a</sup> Important criteria: if 'No' then overall ROB = high; if 'Unclear' then overall ROB = unclear; if 'Yes' then overall ROB = low.  
HRQoL, health-related quality of life.

and the study by Huang *et al.* had an unclear overall RoB judgement, as it was unclear whether missing outcome data were balanced across treatment groups.<sup>87</sup>

### Results of non-randomised studies included in the review

A table of study characteristics and results is presented in [Appendix 6](#). In view of the high RoB of 12 of the 14 included studies – differences in baseline characteristics between treatment groups and treatment allocation being dependent on tumour characteristics for several studies – the results below should be interpreted with caution. The non-randomised nature of these studies and the possibility that some studies may have been undertaken retrospectively mean that these results are less reliable than those of the RCTs described in [Results of RCTs included in the review](#).

### Radiofrequency ablation versus microwave ablation

Two non-randomised studies compared RFA with MWA. One study was assessed as having a low RoB ( $n = 42$  patients).<sup>89</sup> The other study was only reported as a conference abstract and had a high RoB ( $n = 154$  patients).<sup>93</sup> The conference abstract did not report participant inclusion criteria relating to tumour size or the maximum tumour size of the included participants, but the mean size was 2.15 cm in the MWA arm and 1.92 cm in the RFA arm. Patients were unsuitable for percutaneous treatments; the interventions assessed were laparoscopic RFA and laparoscopic MWA.

Only the low-quality study reported OS and disease-free survival rates, which were both higher after laparoscopic RFA than after laparoscopic MWA at 5 years (OS 50% vs. 37%; disease-free survival 19% vs. 12%).<sup>93</sup> However, local tumour progression occurred in more patients in the RFA group than in the MWA group in the low-quality study (21.2% vs. 8.3%) and was similar between groups in the high-quality study (RFA 15% vs. MWA 18.2%). The proportion of patients with a new intrahepatic tumour was also higher in the RFA group than in the MWA group in the high-quality study (20% vs. 4.5%).

Both studies reported that around 95% of patients achieved complete ablation in both arms. After a second treatment, 100% of patients achieved complete ablation in the high-quality study.<sup>89</sup>

The low-quality study reported a similar rate of major complications in both arms (RFA 1% vs. MWA 2%) and no treatment-related deaths in either group.<sup>93</sup> The high-quality study reported only that there were no skin burns, tumour seeding or treatment-related deaths in either group.<sup>89</sup>

### Radiofrequency ablation versus resection

Eight non-randomised studies compared RFA with resection ( $n = 1769$  patients). Seven of the studies had a high RoB<sup>84,85,88,90–92,96</sup> and one had an unclear RoB.<sup>87</sup> Five of the studies with a high RoB were only reported as conference abstracts.<sup>84,85,88,90,91</sup> One study included tumours up to 5 cm in the resection group, but the mean tumour size in this group was 2.1 cm.<sup>96</sup> In four of the studies the treatment received was decided on the basis of patient characteristics (e.g. tumour location) and either there were baseline differences between groups or it was not clearly reported whether this was the case.<sup>85,88,91,96</sup> In one of these studies, group allocation determined which of the two treatments was given as the first choice, but the final decision was based on tumour location.<sup>88</sup> In three of the other studies, allocation to treatment groups was not adequately described and either there were baseline differences between groups or it was unclear whether this was the case.<sup>84,90,92</sup> Only one study reported similar baseline characteristics between treatment groups.<sup>87</sup>

Five of the eight studies reported 1- and 3-year OS rates. In most of these studies, survival was similar between groups at 1 year,<sup>84,85,91,96</sup> although it was slightly higher in the RFA group in one study (RFA 95.9% vs. resection 90.1%).<sup>88</sup> At later time points, findings were more mixed. Three studies reported a higher OS rate in the resection arm at 3 years (RFA vs. resection: 73.9% vs. 83.0%,<sup>84</sup> 74% vs 81%<sup>85</sup> and 76% vs. 83%<sup>91</sup>), but it remained similar in one study (84.5% vs. 84.1%<sup>96</sup>) and was higher in the RFA arm in the other study (75.8% vs. 63.7%<sup>88</sup>). Two studies also reported higher survival rates after RFA at 4 (70.7% vs. 55.5%<sup>88</sup>) or 5 years (50.6% vs. 37.1%<sup>96</sup>).

## RESULTS

Two studies reported recurrence-free survival. In one study it was higher after resection at 1, 3 and 5 years (5-year rate 4.8% vs. 42.9%).<sup>96</sup> In the other study it was similar at 1 year (RFA 74.1% vs. resection 75.9%) but higher after resection at 3 years (40.2% vs. 54.7%).<sup>84</sup>

Findings on recurrence were also mixed. In two studies, recurrence (local and distant/remote) was experienced by more patients in the RFA group (local or distant 85% vs. 42%;<sup>96</sup> local 11.3% vs. 2.0%; remote 53.7% vs. 45.3%<sup>84</sup>). Another two studies reported similar relapse or recurrence rates between groups (1-year relapse rate: RFA 12.9% vs. resection 13.8%;<sup>92</sup> 3-year recurrence rate: RFA 61.7% vs. resection 66%; adjusted HR 0.89, 95% CI 0.72 to 1.1<sup>90</sup>). Two studies reported that no tumours showed local progression or recurrence during the follow-up period in either group.<sup>85,91</sup>

Only two studies reported the complete ablation/resection rate, which was 100% in both treatment arms.<sup>85,91</sup> One study reported quality-of-life outcomes, measured using the FACT-Hep questionnaire.<sup>87</sup> Patients in the RFA group had significantly better HRQoL total scores than those in the resection group after 3, 6, 12, 24 and 36 months.

Data on AEs were limited. Two studies reported a considerably higher rate of total AEs and AEs at grade III or above on the Clavien–Dindo scale after resection.<sup>87,96</sup> While one study reported no hospital deaths in either group,<sup>87</sup> one study reported one hospital death occurring secondary to sepsis in the resection group (RFA 0/40 vs. resection 1/81<sup>96</sup>), and another study reported two cases of treatment-related mortality in the resection group (RFA 0/103 vs. resection 2/92<sup>88</sup>). Two conference abstracts reported only that complication rates were 'comparable' between groups.<sup>85,91</sup>

The average length of hospital stay was approximately twice as long after resection as after RFA in two studies.<sup>87,96</sup> One of the studies also reported that the RFA group experienced a shorter procedure (RFA 44.0 vs. resection 166.5 minutes) and lower blood transfusion rates.<sup>96</sup>

### **Microwave ablation versus resection**

One non-randomised study with a high RoB compared MWA with resection ( $n = 105$  patients).<sup>86</sup> It also included a treatment arm that received PEI. Fewer patients experienced recurrence after MWA than after resection (MWA 38% vs. resection 72%). No data were reported on survival outcomes or AEs.

### **High-intensity focused ultrasound versus radiofrequency ablation**

One non-randomised study with a high RoB compared HIFU with RFA ( $n = 106$  patients).<sup>83</sup> Included patients had primary HCC or first recurrence. Treatment was allocated on the basis of patient characteristics (liver function, decompensated cirrhosis or tumour location), so the groups were not similar at baseline. OS was similar between arms at 1 and 3 years (1 year: HIFU 97.4% vs. RFA 94.6%; 3 years: 81.2% vs. 79.8%). Disease-free survival was also similar at 1 year (HIFU 63.6% vs. RFA 62.4%) but lower in the HIFU group at 3 years (25.9% vs. 34.1%). Complete response was slightly higher in the RFA arm (87.2% vs. 94.9%). More patients in the HIFU group than the RFA group experienced AEs (21.3% vs. 8.5%). However, the rates of AEs at grade III or above on the Clavien–Dindo scale was similar in both groups (HIFU 3/47 vs. RFA 4/59). Patients in the HIFU group had a shorter length of hospital stay than those in the RFA group (median 4 vs. 6 days).

### **Cryoablation versus radiofrequency ablation or microwave ablation**

One non-randomised study with a high RoB compared cryoablation with a group that received either RFA or MWA ( $n = 119$  patients).<sup>95</sup> Results were not reported separately for patients receiving RFA and those receiving MWA. Patients with HCCs up to 5 cm were eligible, but the median tumour size was 2.5 cm in the cryoablation group and 1.9 cm in the RFA/MWA group. Treatment was allocated based on tumour location, so there were baseline differences between the groups.

Overall survival and local recurrence were reported separately for patients with tumours up to 2 cm and patients with tumours over 2 cm. In the  $\leq 2$  cm subgroup, the 2-year OS rate was slightly higher

in the RFA/MWA group (88% vs. 95%) and the 2-year local recurrence rate was similar in both groups (cryoablation 19% vs. RFA/MWA 23%). OS was similar between groups in the > 2 cm subgroup (cryoablation 86% vs. RFA/MWA 85%), but local recurrence occurred in considerably more patients who underwent RFA or MWA than patients who underwent cryoablation (21% vs. 56%).

The 2-year local recurrence-free survival rate (for all tumour sizes) was higher in the cryoablation group (80% vs. 68%). Initial recurrence at other sites of the liver was similar between groups (cryoablation 38% vs. RFA/MWA 34%). Two patients suffered distant metastases in the bone or lung; both were in the cryoablation group.

There was a similar total rate of AEs in the two groups (cryoablation 6/55 vs. RFA/MWA 7/64) and a similar proportion of patients had AEs at grade III or above on the Clavien–Dindo scale (3/55 vs. 3/64). There was no in-hospital mortality in either group. Operative time was longer in the cryoablation group (median 180 vs. 132 minutes). The median length of hospital stay was 8 days in both groups.

### Irreversible electroporation versus radiofrequency ablation

One non-randomised study with a high RoB compared IRE with RFA ( $n = 21$  patients).<sup>94</sup> The maximum tumour size was not reported, but the median size was 2.03 cm in the IRE group and 1.73 cm in the RFA group. Treatment was allocated based on operator preference, tumour size, geometry and location, so there were baseline differences between groups. This study aimed to assess temporal changes in systemic immune responses between these two different types of ablation, and the only relevant data reported were on local tumour progression at 6 months. Local tumour progression was experienced by 1 of 10 patients in the IRE group and 0 of 11 patients in the RFA group.

### Ongoing trials

The electronic searches for non-randomised trials identified two potentially relevant ongoing RCTs that were not identified in the RCT searches (described in [Ongoing trials](#)). Further details are presented in [Table 13](#).

## Updated network meta-analyses using RCT and non-RCT evidence

Of the 14 non-randomised studies that were included in the systematic review, two<sup>87,89</sup> could be included in the updated NMAs. Huang (2014) reported data that could be incorporated in the NMAs for OS and PFS,<sup>87</sup> while Qian (2012) reported data that could be incorporated in the NMA for PFS.<sup>89</sup>

Data from a prospective registry of patients undergoing treatment for HCC were made available to the research team by a research group at Leeds Teaching Hospitals NHS Trust (Dr Tze Wah, Leeds Teaching Hospitals NHS Trust, 5 October 2021, personal communication). This contained data for 303 patients who had received either RFA, MWA, IRE or cryoablation for primary HCC. Most patients received RFA, with a smaller number receiving MWA. Very few patients received IRE or cryoablation. Data were unpublished at the time of our analysis, but have been submitted for publication.

**TABLE 13** Table of potentially relevant ongoing RCTs (identified from non-RCT searches)

Study	Further details
ChiCTR2000039404	Clinical trial register record describing a single-centre RCT comparing SBRT vs. RFA for $\leq 2$ cm small HCC. Registered 2020.
ClinicalTrials.gov: NCT04891874	Clinical trial register record describing a single-centre RCT comparing adjuvant SBRT after surgery vs. surgery alone for early-stage HCC with microvascular invasion and narrow resection margin. Sponsor: Eastern Hepatobiliary Surgery Hospital, China. Trial status: Completed, last update posted 10 September 2021.

Data from the Leeds patients were reported for numerous outcomes. There were sufficient data for inclusion in NMAs for OS, PFS, and local recurrence.

As there was no new evidence for overall recurrence, no updated NMAs or threshold analyses were conducted for this outcome.

### Overall survival

#### Data

The network diagram for OS is presented in [Figure 19](#). In addition to the randomised studies included in the NMA in [Overall survival](#), one two-arm and one three-arm study provided non-randomised evidence for one new intervention in addition to three interventions already included in the network. A summary of the additional non-randomised evidence included in the NMA is provided in [Report Supplementary Material 4](#).

#### Model selection and inconsistency checking

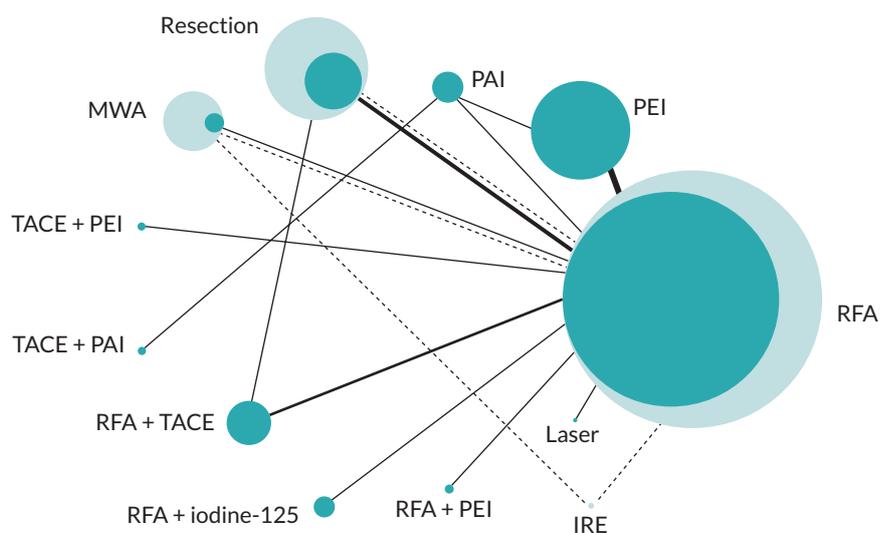
Model fit parameters for the FE and RE models are presented in [Report Supplementary Material 5](#). All three models fit the data well, but as the difference in the DICs between the FE and RE models was  $< 3$ , the simpler FE model was chosen.

The 95% CrI for the model using the half-normal  $(0, 0.50^2)$  prior was wider than the 95% CrI for the model using the half-normal  $(0, 0.19^2)$ . This shows that the estimate of between-study heterogeneity is sensitive to the level of prior heterogeneity assumed due to few studies being included for each comparison in the network. Plots for the prior and posterior distributions of the between-study heterogeneity for the RE models are presented in [Report Supplementary Material 5](#). There was no evidence to suggest inconsistency in the network. Details of the inconsistency check and node-splitting results are presented in [Report Supplementary Material 5](#).

#### Model results

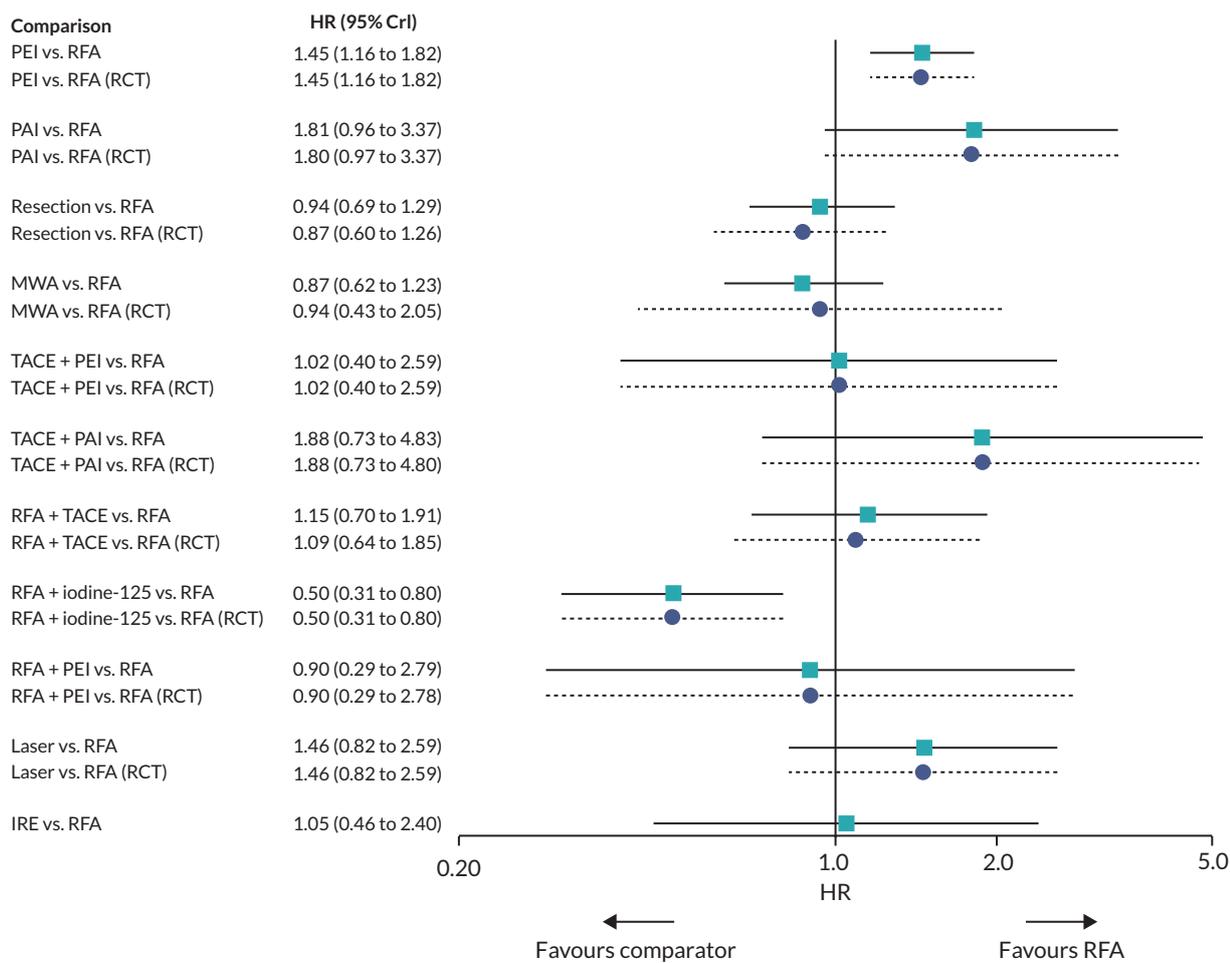
Hazard ratios for OS for all treatments compared with RFA are presented in [Figure 20](#).

The results for the NMA were not very different from the results from the NMA comparing only randomised evidence (see [Model results](#)). With the addition of non-randomised studies there was also



**FIGURE 19** Network diagram for OS.

Treatment nodes in the network diagram are scaled proportional to the number of patients who receive a particular treatment. The widths of the lines joining two nodes are weighted according to the number of studies that provide evidence on that comparison. The light blue circles represent the number of patients who receive a particular treatment in both randomised and non-randomised studies, and dashed lines represent comparisons that are added to the network by non-randomised evidence.



**FIGURE 20** Plot of HRs for OS compared with RFA for the FE model. HRs < 1 favour the comparator treatment over RFA.

evidence to suggest that RFA + iodine-125 improves survival compared with resection. There was also evidence to suggest that MWA improves survival compared with PEI and PAI. HRs comparing all treatment groups against each other for FE and RE models are reported in [Report Supplementary Material 5](#).

The mean and median ranks for each treatment, with their corresponding 95% CrIs, are presented in [Table 14](#). RFA + iodine-125 had the highest probability of being ranked the best treatment. However, as seen for NMAs including only randomised evidence (see [Model results](#)), there was a high level of uncertainty in treatment rankings, also visible in the treatment rank plots (see [Figure 21](#)).

### Progression-free survival

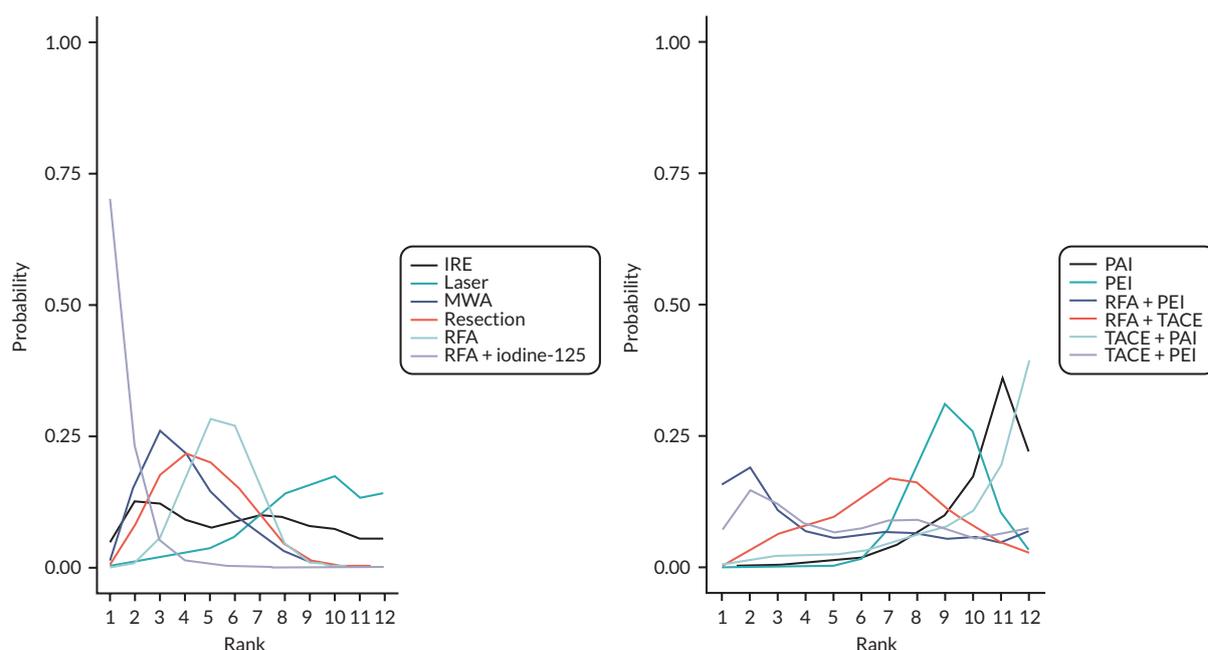
#### Data

The network diagram for PFS is presented in [Figure 22](#). In addition to the randomised studies included in the NMA in [Progression-free survival](#), two two-arm and one three-arm study provided non-randomised evidence for two new interventions in addition to two interventions already included in the network. A summary of the additional non-randomised evidence included in the NMA is provided in [Report Supplementary Material 4](#).

## RESULTS

**TABLE 14** Mean and median treatment ranks for the FE model, with corresponding 95% CrIs for OS, sorted by mean rank out of 12 treatments

Treatments	Mean rank	Median rank	95% CrI of the rank
RFA + iodine-125	1.40	1	(1.00 to 3.00)
MWA	4.10	4	(2.00 to 8.00)
Resection	4.76	5	(2.00 to 8.00)
RFA + PEI	5.19	4	(1.00 to 12.00)
RFA	5.46	5	(3.00 to 8.00)
TACE + PEI	5.88	6	(1.00 to 12.00)
IRE	6.08	6	(1.00 to 12.00)
RFA + TACE	6.94	7	(2.00 to 12.00)
Laser	8.86	9	(3.00 to 12.00)
PEI	9.14	9	(7.00 to 12.00)
TACE + PAI	9.95	11	(3.00 to 12.00)
PAI	10.23	11	(5.00 to 12.00)

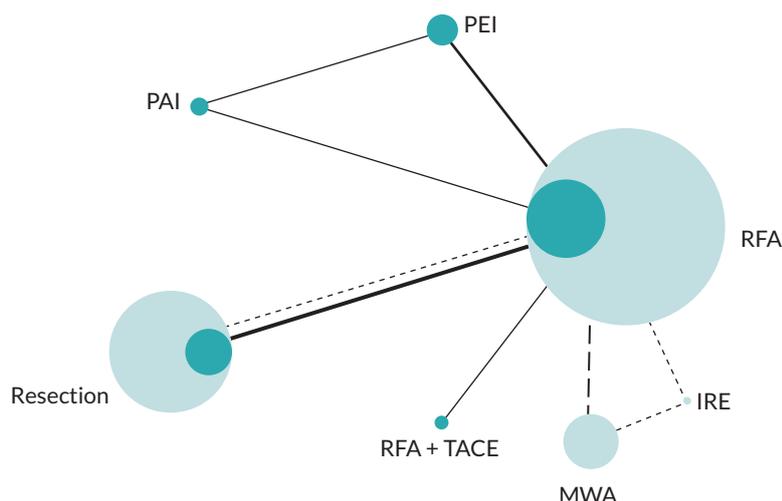


**FIGURE 21** Rank plot for OS for the FE model

### Model selection and consistency checking

Model fit parameters for the FE and RE models are presented in [Report Supplementary Material 5](#). All three models fit the data well, but as the difference in the DICs between the FE and RE models was  $< 3$ , the simpler FE model was chosen.

The between-study heterogeneity was low for the two RE models. However, the 95% CrI for the model using the half-normal  $(0, 0.50^2)$  prior was wider than the 95% CrI for the model using the half-normal  $(0, 0.19^2)$  indicating that the estimate of between-study heterogeneity is sensitive to the level of prior heterogeneity assumed due to few studies being included for each comparison in the network.



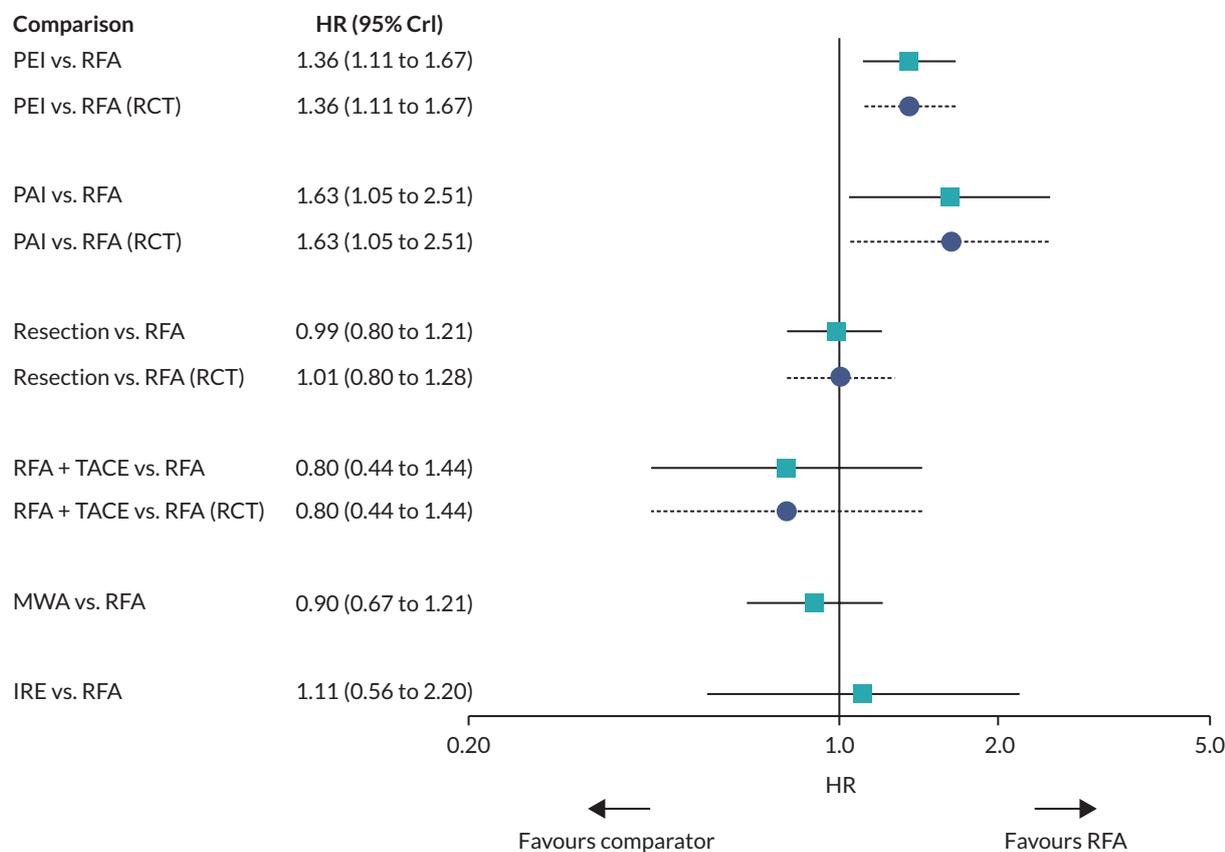
**FIGURE 22** Network diagram for PFS.

Treatment nodes in the network diagram are scaled proportional to the number of patients who receive a particular treatment. The widths of the lines joining two nodes are weighted according to the number of studies that provide evidence on that comparison. The light blue circles represent the number of patients who receive a particular treatment in both randomised and non-randomised studies, and dashed lines represent comparisons that are added to the network by non-randomised evidence.

There is no potential for inconsistency in this network as there is no independent, indirect evidence for any of the comparisons – the two loops are formed by two three-arm studies, one of which is the work at Leeds Teaching Hospitals NHS Trust described above (Dr Tze Wah, personal communication).<sup>53</sup>

### Model results

HRs for PFS for all treatments compared with RFA are presented in [Figure 23](#).



**FIGURE 23** Plot of HRs for PFS compared with RFA for the FE model.

HRs < 1 favour the comparator treatment over RFA.

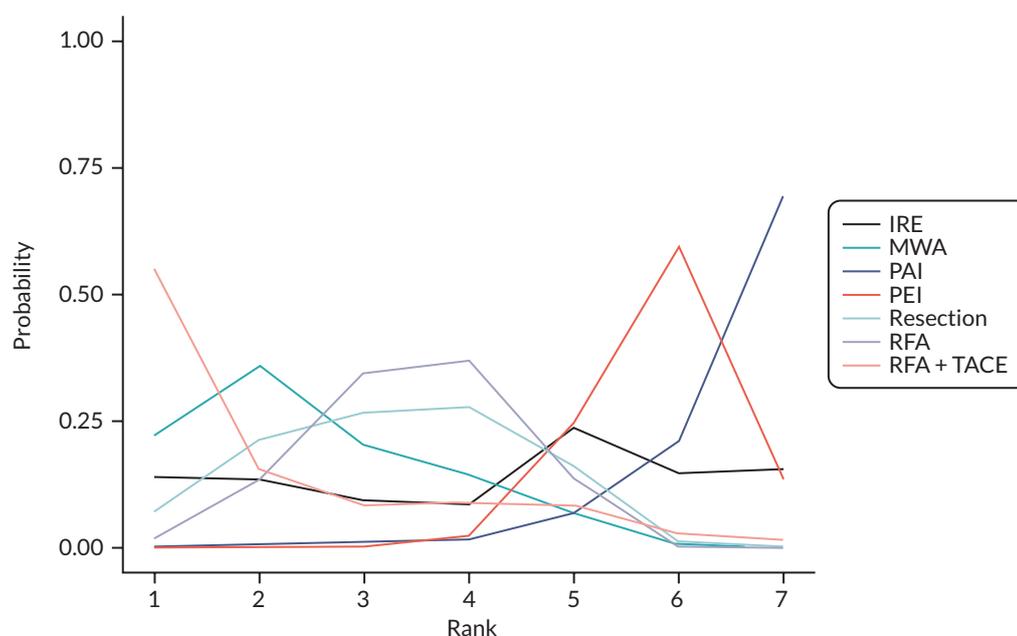
Similar to the NMA using only RCT evidence (see *Model selection and consistency checking*), there was evidence to suggest that PEI and PAI worsen PFS compared with RFA. However, with the addition of the non-randomised studies, there was also evidence to suggest that resection and MWA improved PFS compared with PEI and PAI. HRs comparing all treatment groups against each other for FE and RE models are reported in [Report Supplementary Material 5](#).

The treatment rank plot for PFS is presented in [Figure 24](#), and the mean and median ranks for each treatment, with their corresponding 95% CrIs, are presented in [Table 15](#). RFA + TACE had the highest probability to be ranked the best treatment. However, there was a high level of uncertainty in the treatment ranking – all treatments displayed wide CrIs for ranks.

### Local recurrence

#### Data

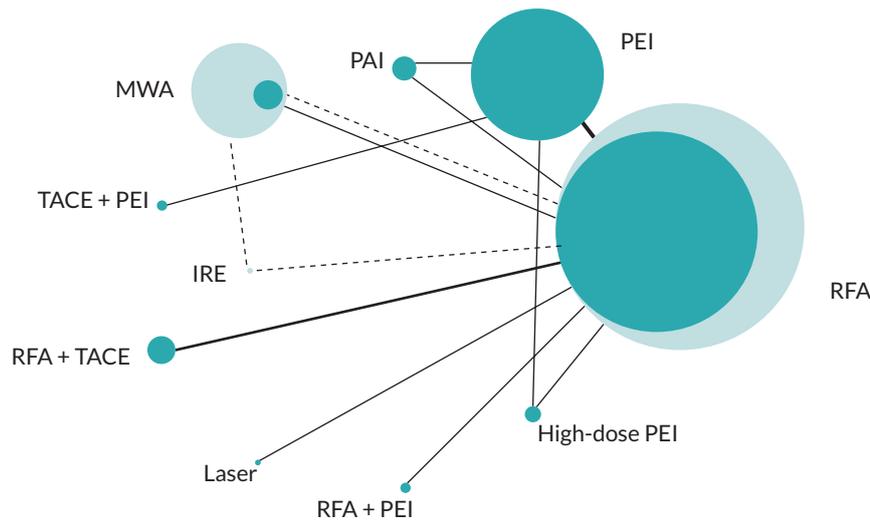
The network diagram for local recurrence is presented in [Figure 25](#). In addition to the randomised studies included in the NMA in *Overall recurrence*, one three-arm study provided non-randomised evidence for one new intervention in addition to two interventions already included in the network. A summary of the additional non-randomised evidence included in the NMA is provided in [Report Supplementary Material 4](#).



**FIGURE 24** Rank plot for PFS for the FE model.

**TABLE 15** Mean and median ranks, with corresponding 95% CrIs for PFS for the FE model, sorted by mean rank

Treatments	Mean rank	Median rank	95% CrI
RFA + TACE	2.14	1	(1.00 to 6.00)
MWA	2.50	2	(1.00 to 5.00)
Resection	3.28	3	(1.00 to 5.00)
RFA	3.48	4	(2.00 to 5.00)
IRE	4.21	5	(1.00 to 7.00)
PEI	5.83	6	(4.00 to 7.00)
PAI	6.55	7	(4.00 to 7.00)



**FIGURE 25** Network diagram for local recurrence.

Treatment nodes in the network diagram are scaled proportional to the number of patients who receive a particular treatment. The widths of the lines joining two nodes are weighted according to the number of studies that provide evidence on that comparison. The light blue circles represent the number of patients who receive a particular treatment in both randomised and non-randomised studies, and dashed lines represent comparisons that are added to the network by non-randomised evidence.

### Model selection and inconsistency checking

Model fit parameters for the three models are reported in [Report Supplementary Material 5](#). All three models fit the data well. The between-study heterogeneity was low and consistent for the two RE models. However, the 95% CrI for the model using the half-normal (0, 0.50<sup>2</sup>) prior was wider than the 95% CrI for the model using the half-normal (0, 0.19<sup>2</sup>).

As the difference in the DICs between the FE and RE models was < 3, the simpler FE model was chosen. Plots for the prior and posterior distributions of the between-study heterogeneity for the RE models are presented in [Report Supplementary Material 5](#).

There is no potential for inconsistency in this network as there is no independent, indirect evidence for any of the comparisons; the three loops in the network are formed by three separate three-arm studies.

### Model results

Relative risks for local recurrence for all treatments compared with RFA are presented in [Figure 26](#).

Similar to the NMA using only RCT evidence (see [Model results](#)), there was evidence to suggest that PEI increased the risk of local recurrence compared with RFA, and that RFA + PEI decreased the risk of local recurrence compared with PEI. However, with the addition of non-randomised studies, there was also now evidence to suggest that IRE increased the risk of local recurrence compared with RFA and RFA + PEI, although the CrIs for both comparisons were very wide. RRs comparing all treatment groups against each other for FE and RE models are reported in [Report Supplementary Material 5](#).

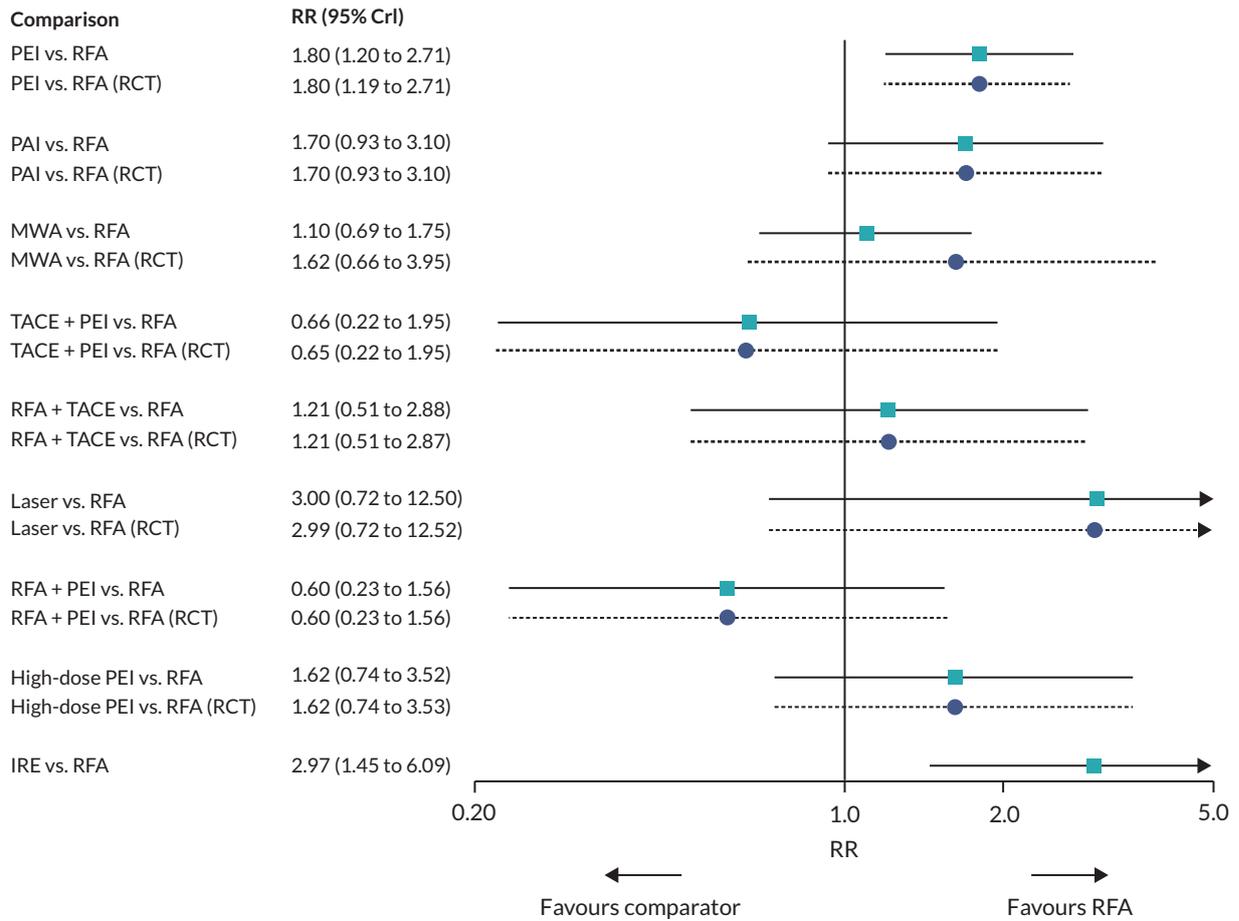
The treatment rank plot for local recurrence is presented in [Figure 27](#), and the mean and median ranks for each treatment, with their corresponding 95% CrIs, are presented in [Table 16](#). There was a high level of uncertainty in treatment ranks; all treatments had rank probabilities below 50% for all treatment ranks.

## Updated threshold analysis

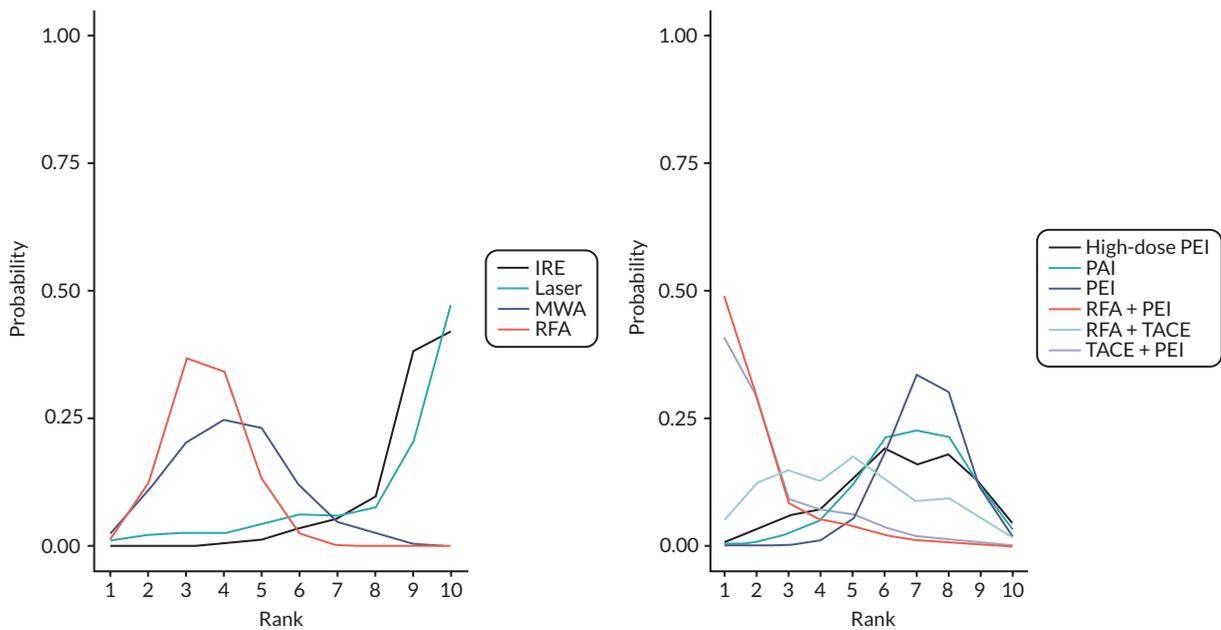
### Overall survival

The forest plot for the threshold analysis is presented in [Figure 28](#).

## RESULTS



**FIGURE 26** Plot of RRs for local recurrence compared with RFA for the FE model. RRs < 1 favour the comparator treatment over RFA.



**FIGURE 27** Rank plot for local recurrence for the FE model.

**TABLE 16** Mean and median ranks, with corresponding 95% CrIs, for local recurrence, sorted by mean rank, for the FE model

Treatments	Mean rank	Median rank	95% CrI
RFA + PEI	2.03	2	(1.00 to 6.00)
TACE + PEI	2.38	2	(1.00 to 7.00)
RFA	3.53	3	(2.00 to 5.00)
MWA	4.23	4	(2.00 to 8.00)
RFA + TACE	4.90	5	(1.00 to 9.00)
High-dose PEI	6.45	7	(2.00 to 10.00)
PAI	6.78	7	(3.00 to 10.00)
PEI	7.25	7	(5.00 to 9.00)
Laser	8.40	9	(2.00 to 10.00)
IRE	9.05	9	(6.00 to 10.00)

Interventions that included PEI and PAI were not considered in the threshold analysis, and therefore comparisons including those interventions – PEI versus RFA (2 vs. 1), PAI versus RFA (3 vs. 1), RFA + PEI versus RFA (10 vs. 1), PAI versus PEI (3 vs. 2), TACE + PEI versus PEI (6 vs. 2), and TACE + PAI versus PAI (7 vs. 3) – had very large thresholds on the log scale. The following comparisons also had very large thresholds on the log scale: RFA + TACE versus RFA (8 vs. 1), IRE versus RFA (11 vs. 1). None of the other comparisons have thresholds that indicate estimates are sensitive to small changes in log-HRs. The thresholds and new optimum treatments, based only on relative effects, are presented in [Report Supplementary Material 5](#).

### Progression-free survival

The forest plot for the threshold analysis is presented in [Figure 29](#).

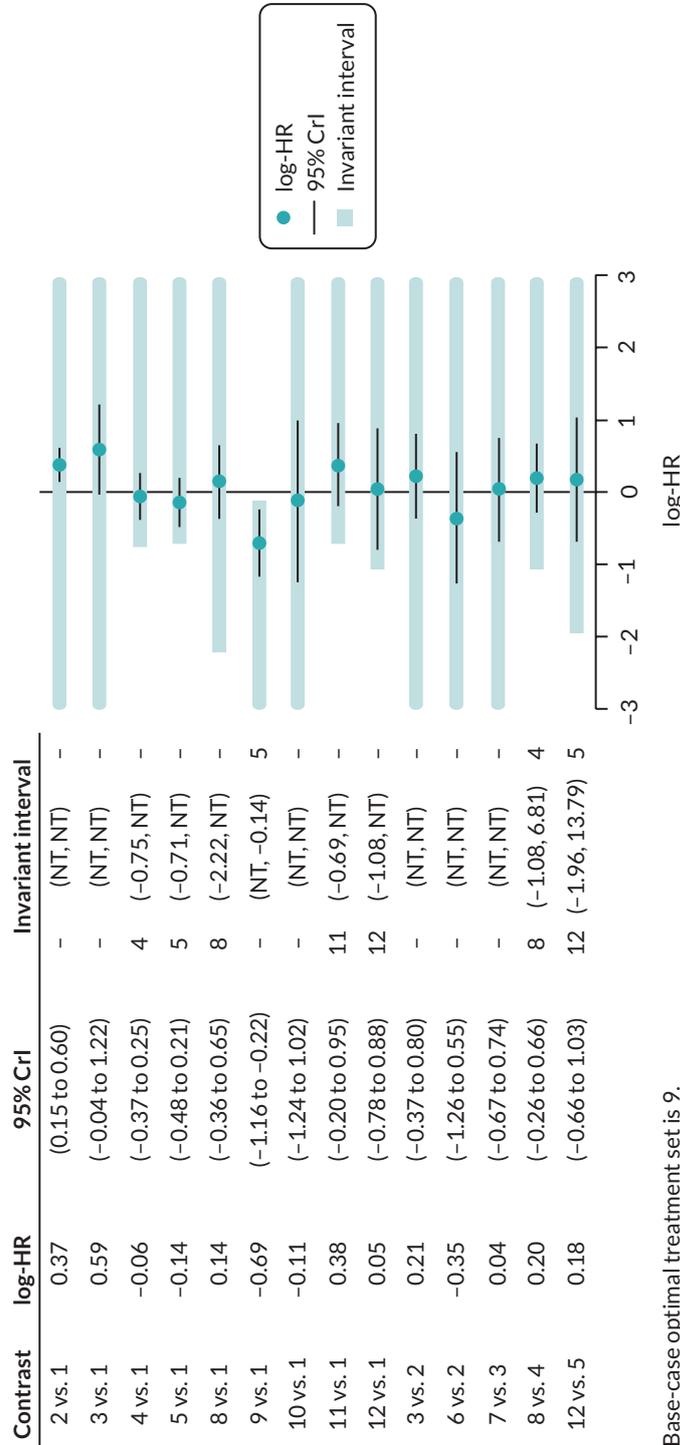
Credible intervals for the RFA + TACE versus RFA (5 vs. 1), MWA versus RFA (6 vs. 1), and IRE versus RFA (7 vs. 1) comparisons extend beyond the limits of the invariance intervals, suggesting that the recommended treatment is sensitive to the uncertainty in the data, changing the optimum treatment to MWA for the RFA + TACE versus RFA and MWA versus RFA comparisons and to IRE for the IRE versus RFA comparison.

Three comparisons that included PEI and PAI – PEI versus RFA (2 vs. 1), PAI versus RFA (3 vs. 1), and PAI versus PEI (3 vs. 2) – had very large thresholds on the log scale. The negative threshold for the resection versus RFA comparison (4 vs. 1) was very small, and a change of 0.21 units on the log-HR scale in the negative direction changes the optimum treatment to resection. The positive threshold for the RFA + TACE versus RFA comparison (5 vs. 1) was very small, and a change of 0.12 units in the positive direction changes the optimum treatment to MWA. Similarly, the negative threshold for the MWA versus RFA comparison (6 vs. 1) was very small, and a change of 0.13 units in the negative direction changes the optimum treatment to MWA.

Thresholds and new optimum treatments, based only on relative effects, are presented in [Report Supplementary Material 5](#).

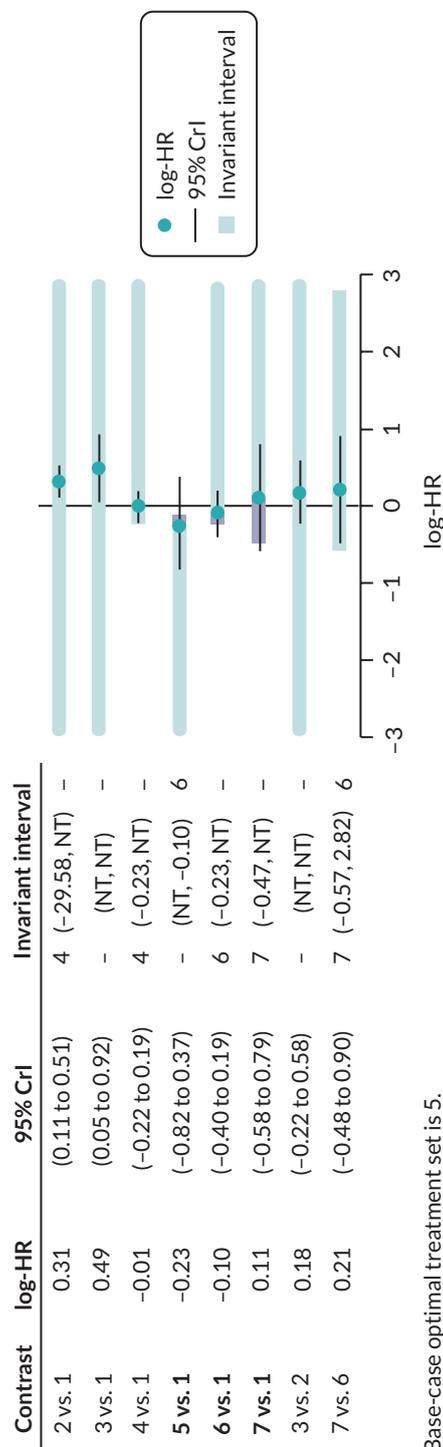
### Local recurrence

The forest plot for the threshold analysis is presented in [Figure 30](#).

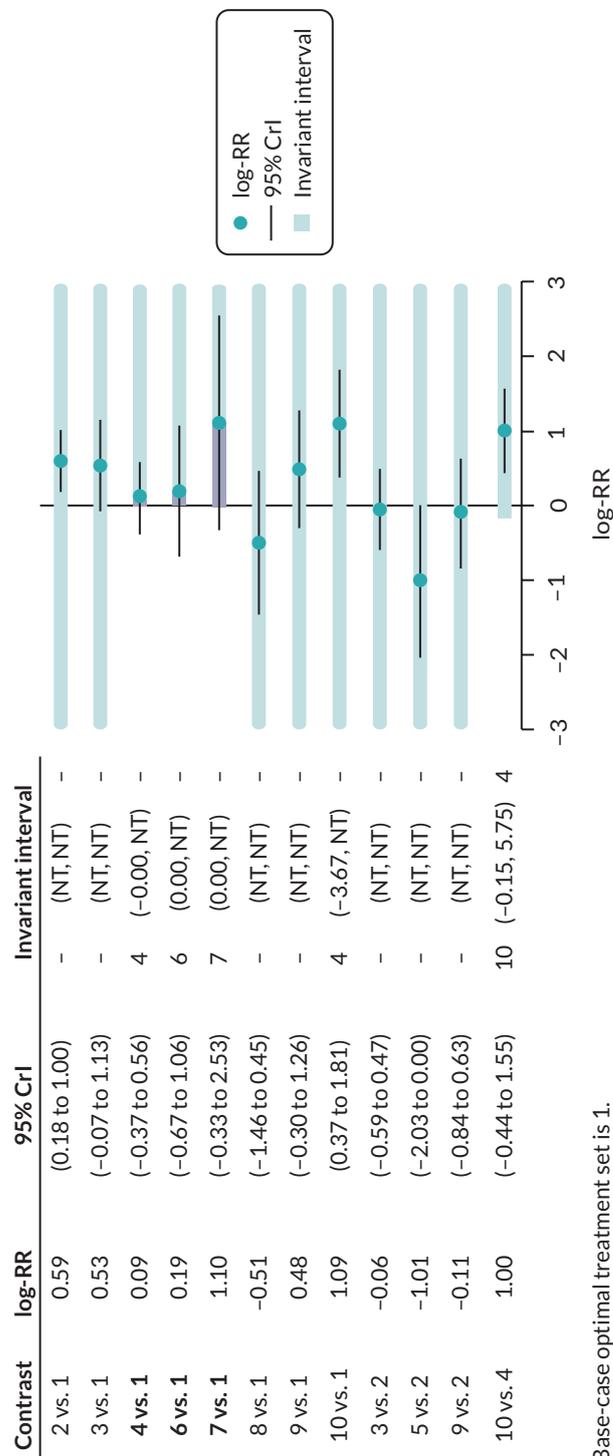


Base-case optimal treatment set is 9.

**FIGURE 28** Forest plot for threshold analysis results for OS for the updated NMA. Treatment codes: 1: RFA, 2: PEI, 3: PAI, 4: Resection, 5: MWA, 6: TACE + PEI, 7: TACE + PAI, 8: RFA + TACE, 9: RFA + iodine-125, 10: RFA + PEI, 11: Laser, 12: IRE. The optimum treatment for this analysis was RFA + iodine-125.



**FIGURE 29** Forest plot for the threshold analysis for PFS for the updated NMA. The light blue shaded bars represent the invariant interval; where the bars are shown in purple, CrIs for the comparison extend beyond the invariant interval. Treatment codes: 1: RFA, 2: PEI, 3: PAI, 4: Resection, 5: RFA + TACE, 6: MWA, 7: IRE. The optimum treatment for this analysis was RFA + TACE.



Base-case optimal treatment set is 1.

**FIGURE 30** Forest plot for results of threshold analysis for local recurrence for the updated NMA.

The light blue shaded bars represent the invariant interval; where the bars are shown in purple, CrIs for the comparison extend beyond the invariant interval.

**Treatment codes:** 1: RFA, 2: PEI, 3: PAI, 4: MWA, 5: TACE + PEI, 6: RFA + TACE, 7: Laser, 8: RFA + PEI, 9: High-dose PEI, 10: IRE. The optimum treatment for this analysis was RFA + PEI.

Credible intervals for the MWA versus RFA (4 vs. 1), RFA + TACE versus RFA (6 vs. 1), and laser versus RFA (7 vs. 1) comparisons extend beyond the limits of the invariance intervals, suggesting that the recommended treatment is sensitive to the uncertainty in the data, changing the optimum treatment to MWA, RFA + TACE and laser, respectively.

Seven comparisons that included PEI and PAI – PEI versus RFA (2 vs. 1), PAI versus RFA (3 vs. 1), RFA + PEI versus RFA (8 vs. 1), high-dose PEI versus RFA (9 vs. 1), PAI versus PEI (3 vs. 2), TACE + PEI versus PEI (5 vs. 2), high-dose PEI versus PEI (9 vs. 2) – had very large thresholds on the log scale.

The negative thresholds for the MWA versus RFA (4 vs. 1) and RFA + TACE versus RFA (6 vs. 1) comparisons were very small, and changes of 0.09 and 0.19 units in the negative direction change the optimum treatments to MWA and RFA + TACE, respectively.

Thresholds and new optimum treatments, based only on relative effects, are presented in [Report Supplementary Material 5](#).



## Chapter 5 Feasibility of economic modelling

This section considers the feasibility of developing a de novo economic model to inform a cost-effectiveness and value of information (VOI) analysis considering ablative and non-surgical therapies for the treatment of small HCC tumours. In considering the feasibility of an appropriate economic evaluation, it is assumed that the developed model will be consistent with the NICE reference case,<sup>97</sup> adopting a UK perspective and using a cost-utility approach accounting for both the relevant costs and benefits of the assessed technology.

### Approach

Assessment of the feasibility of undertaking economic evaluation and VOI analysis was considered by conducting a targeted review exploring previous economic analyses evaluating technologies for the treatment of HCC; see *Review methods* below for details of methods used. Studies identified in the review were then summarised to consider key features and what data are typically required to support these models. Based on these previous evaluations and in consultation with clinical experts, a conceptual model was then developed to consider an appropriate model structure that could be used in any future economic analysis.

The availability of data to inform an economic analysis was considered. This assessment covered the availability of relevant clinical evidence (based principally on the clinical effectiveness review). The availability of evidence concerning quality of life, resource use and costs was also considered; this was informed by evidence identified as part of the clinical effectiveness review, the identified cost-effectiveness studies and established sources of relevant data.

### Cost-effectiveness review

#### Review methods

Targeted literature searches were adapted from the search strategies used to identify RCTs (see [Appendix 1](#)) and included terms for small or early HCC and a broad set of terms aimed at identifying any economic evidence. The following databases were searched in May 2021:

- Ovid MEDLINE(R) ALL: 1946 to 12 May 2021
- Embase: 1974 to 12 May 2021
- NHS Economic Evaluation Database
- Econlit: 1886 to 29 April 2021.

Study design search filters for economic papers were applied to Ovid MEDLINE and Ovid Embase only. The *Canadian Journal of Health Technologies* (CADTH)'s<sup>98</sup> narrow economic filter was used on MEDLINE and was adapted for use on Embase. No language or geographical restrictions were applied to the searches across any of the databases. A date limit of 2000 onwards was applied to the searches to align with the clinical effectiveness review. Details of the search strategies used are reported in [Report Supplementary Material 1](#).

Study selection was conducted in two stages: (1) titles and abstracts were examined and screened for any study potentially relevant to the cost-effectiveness review; and (2) full texts were then obtained and screened for inclusion. A single reviewer screened all studies.

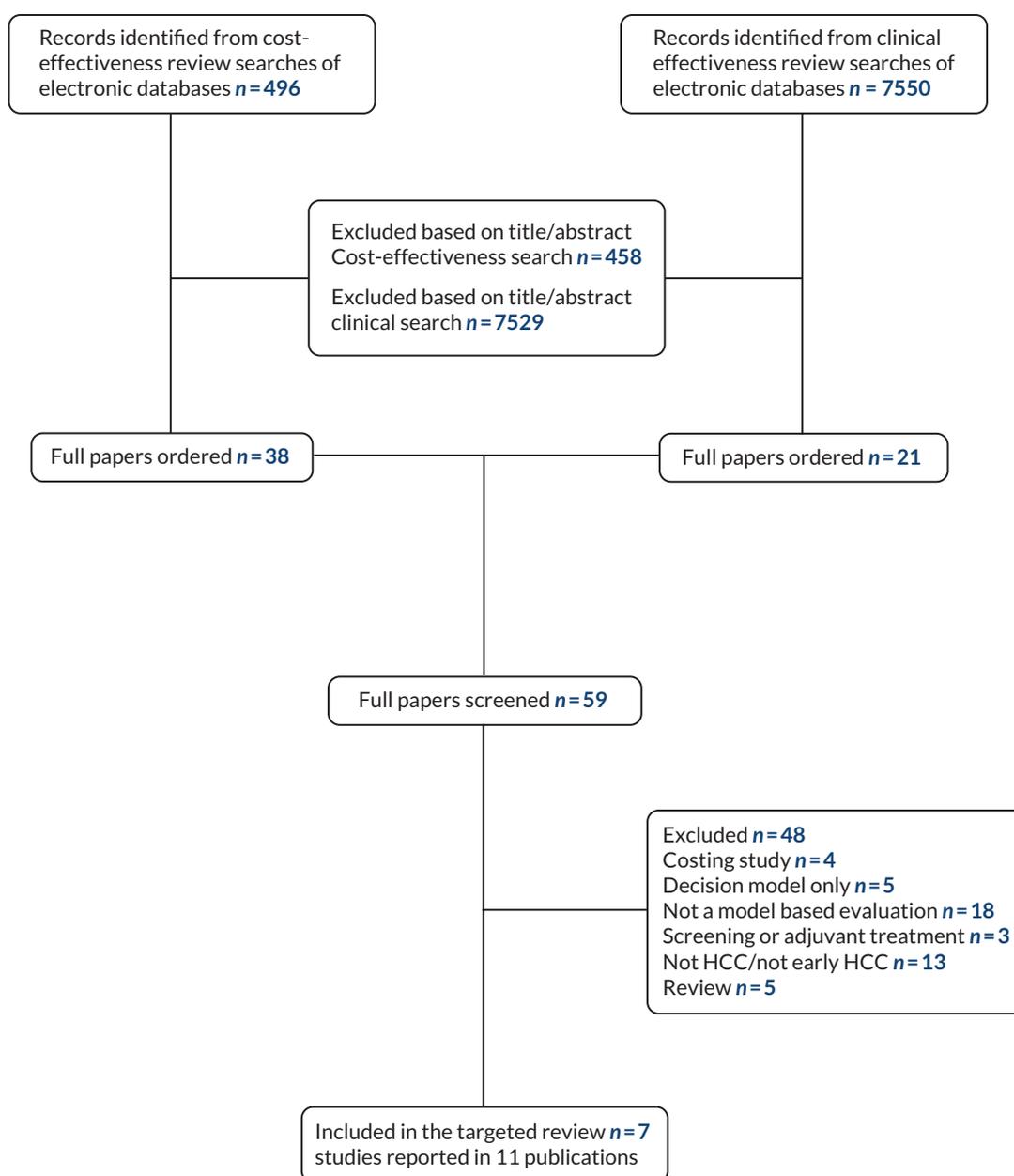
Studies were included in the review if they assessed the cost-effectiveness of any technology for the treatment of very early/early HCC; note that this is broader than the inclusion criteria for the clinical effectiveness review. A broad range of studies was considered for inclusion in the review, including

economic evaluations conducted alongside trials, modelling studies, and analyses of administrative databases. Only full economic evaluations comparing two or more options and including both costs and consequences (cost-effectiveness, cost-utility or cost-benefit analyses) were included.

Studies meeting the inclusion criteria were summarised, noting key features including the model structure adopted, key assumptions and any data reported that may be relevant to undertaking an economic evaluation of ablative and non-surgical therapies for early HCC. As this was not intended to be a formal review of cost-effectiveness studies, study quality was not assessed.

## Results

A flow diagram describing study selection is presented in [Figure 31](#). Searches of the literature for economic evidence identified 496 papers following the removal of duplicates, with 38 identified for full



**FIGURE 31** Flow diagram of the study selection process (cost-effectiveness review).

text review. A further 21 papers were identified for full text review as part of the clinical effectiveness review, making a total of 59 papers. Following the selection process, seven studies reported in 11 publications were found to meet the eligibility criteria and were included in the review.

An overview of the study characteristics for each included study is presented in [Table 17](#). The majority of studies evaluated two treatment alternatives. Interventions evaluated included liver transplant, resection, RFA, SIRT, and TACE. In UK clinical practice, the use of SIRT, TACE and liver transplant for the treatment of small HCCs is limited; their inclusion in the identified studies reflects the broad inclusion criteria and national differences in clinical practice.

**TABLE 17** Data extraction: cost-effectiveness review

**Cucchetti (2013)**<sup>99-101</sup>

Model structure	The modelling approach is not fully clear; described as a Markov model, but potentially adopts a semi-Markov or simulation approach. Model considers survival, recurrence and Child-Pugh status.
Time horizon, perspective and discounting	Time horizon was not stated. Perspective and setting were not stated, though the majority of the costs were drawn from Italian national health system. Costs and benefits were discounted at a rate of 3%.
Population	Patients within the Milan criteria up to three tumours < 3 cm, or one tumour up to 5 cm.
Intervention and comparators	Resection vs. RFA.
Clinical evidence	Parametric extrapolation of survival data (OS and disease-free survival) appears to have been undertaken though the specific approach adopted is unclear. Hazard rates applied to model treatment efficacy were based on the proportion of patients achieving 3-year survival/3-year disease-free survival and were drawn from a meta-analysis of relevant studies. The model also drew on evidence of hospital length of stay which was drawn from the meta-analysis and parameterised in the model.
HRQoL	Health state utilities were based on values reported in the literature, including a review by McLemon <i>et al.</i> <sup>110</sup> Values did not vary by treatment received and were not specific to HCC.
Resources and costs	Cost categories modelled included procedure costs, length of stay, costs of subsequent treatments and patient follow-up costs. Costs applied in the model were obtained from Medicare and Italian national health system sources.

**Lai (2014)**<sup>103</sup>

Model structure	Markov model with the following health states: small HCC < 3 cm tumour, cancer-free, progressive HCC and death. Additional tunnel states were also used to count the number of ablation procedures, with a maximum of three permitted.
Time horizon, perspective and discounting	Time horizon appeared to be lifetime horizon (until 99% of patients were dead). A Chinese health-care setting was considered, but the perspective was not stated formally. Costs and benefits were discounted at a rate of 3%.
Population	Patients with a solitary, small tumour < 3 cm and Child-Pugh class A or B.
Intervention and comparators	Real-time virtual sonography-guided ablation vs ultrasound-guided ablation.
Clinical evidence	Probabilities for each outcome were drawn from the literature, with the majority of inputs drawn from Cho <i>et al.</i> <sup>111</sup> Efficacy was not determined using comparator estimates of effect. Outcomes considered included mortality rates (with separate rates applied to cirrhotic patients, tumour-free patients, progressed HCC), ablation success rate, rate of local recurrence, distant recurrence, probability of seeding tumour (RFA only), liver transplant rate, procedure-related mortality and procedure-related complications.
HRQoL	Health state utilities were based on values reported in the literature, including McLemon <i>et al.</i> <sup>110</sup> Values did not vary by treatment received and were not specific to HCC.
Resources and costs	Cost categories considered included procedure costs, inpatient administration costs associated with RFA, disease management and follow-up care costs, terminal care costs, and AE costs. Values were drawn from the literature and did not consider any UK relevant sources.

continued

**TABLE 17** Data extraction: cost-effectiveness review (*continued*)**Lim (2015)<sup>109</sup>**

Model structure	Markov cohort model with alternative model structures applied according to treatment received. In the liver resection arm the following health states were modelled: compensated cirrhosis, decompensated cirrhosis, HCC recurrence, dead. In the liver transplant arm the following states were modelled: waiting list compensated cirrhosis, waiting list decompensated cirrhosis, liver transplant contraindicated, post liver transplant, dead.
Time horizon, perspective and discounting	Time horizon was not reported; a payer perspective was adopted though setting was not clear. Costs and benefits discounted at a rate of 3%.
Population	Patients within the Milan criteria up to three tumours < 3 cm, or one tumour up to 5 cm.
Intervention and comparators	Liver resection vs. liver transplant.
Clinical evidence	Evidence was drawn from multiple sources identified in the literature and did not rely on comparative assessment of effectiveness. Outcomes modelled included: decompensation risk, decompensated cirrhosis-related survival, postoperative risks (liver resection and liver transplant), post liver resection recurrence rate, wait list time, dropout risk and survival.
HRQoL	Health state utilities were based on values reported in the literature, though the specific studies used were not reported. Values did not vary by treatment received.
Resources and costs	Cost categories considered included: procedure costs, and disease management and follow-up costs. Costs were drawn from a systematic review of values reported in the literature and the median reported value used. Where data were unavailable, clinical expert opinion was used. Costs used were not directly relevant to the UK.

**Naugler (2010)<sup>106</sup>**

Model structure	Markov model using two distinct structures for each arm. In the watchful waiting arm the following health states were modelled: monitoring without therapy, tumour progression inside Milan criteria, tumour progression outside Milan criteria, liver decompensation, and death. In the immediate treatment arm the following health states were modelled: HCC therapy, tumour progression inside Milan criteria, liver decompensation, and death.
Time horizon, perspective and discounting	Time horizon was 10 years. Perspective and setting were not stated; costs were however, drawn from the US health system. Costs and benefits were discounted at a rate of 3%.
Population	Patients with tumours < 2 cm, not eligible for resection but eligible for transplant with compensated cirrhosis.
Intervention and comparators	Watchful waiting vs. immediate treatment with TACE vs. immediate treatment with RFA.
Clinical evidence	Probabilities for each outcome were drawn from the literature using multiple sources. Efficacy was not determined using comparator estimates of effect. Outcomes modelled in the watchful waiting arm included: tumour progression inside/outside Milan criteria, and survival inside/outside Milan criteria. Outcomes modelled in the immediate treatment arm included: survival within Milan criteria, survival without progression, tumour progression inside Milan criteria. In both arms the model also considered liver decompensation risk, liver transplant rate, and post-transplant survival.
HRQoL	Not considered.
Resources and costs	Cost categories modelled included procedure costs, disease management and patient follow-up costs, drug acquisition costs. Costs applied in the model were obtained from Medicare and were not relevant to a UK setting.

**Rostambeigi (2014)<sup>107,108</sup>**

Model structure	The model used a simulation approach. The structure adopted was not clearly reported, but appeared to allow for disease recurrence, mortality, and liver transplant.
Time horizon, perspective and discounting	Time horizon was not stated. Perspective and setting were not stated; costs were however, drawn from the US health system. Discounting of future costs and benefits does not appear to have been applied.
Population	BCLC A.

**TABLE 17** Data extraction: cost-effectiveness review (*continued*)

Intervention and comparators	SIRT vs. TACE.
Clinical evidence	Probabilities for each outcome were drawn from exponential curves and used to estimate survival based on reported survival rates. Other outcomes considered include recurrence and re-treatment of HCC, and transplant rates.
HRQoL	Not considered.
Resources and costs	Cost categories modelled included procedure costs, AEs, and patient follow-up costs. Costs applied in the model were obtained from Medicare reimbursement costs and were not directly relevant to the UK.

**Sarasin (2001)**<sup>102</sup>

Model structure	A Markov model was developed that accounted for wait time for transplant. Modelled health states included: cirrhosis, HCC, no contraindications to CLT, cured HCC and cirrhosis, contraindications to CLT/palliative care, and death.
Time horizon, perspective and discounting	Time horizon was not stated. A US payer perspective was adopted using 1998 prices. Costs and benefits were discounted at a rate of 3%.
Population	Early HCC – single HCC not exceeding 5 cm in diameter, or up to three tumours up to 3 cm in size, in the absence of vascular or extrahepatic involvement.
Intervention and comparators	CLT vs. LDLT.
Clinical evidence	Parameter inputs were identified via searches of the literature. Outcomes were determined by wait time (2 months for LDLT, 6 months for CLT), probability of developing contraindications, donor mortality, palliative care mortality and post-transplant mortality. Transplant outcomes between CLT and LDLT were assumed to be the same.
HRQoL	Utility values were informed by the literature and did not vary by treatment received.
Resources and costs	Cost categories modelled included chemoembolisation costs incurred while waiting for transplant, transplant-related costs (assumed to be the same for CLT and LDLT), donor assessment (accounting for failures to proceed), and disease management and patient follow-up costs. Costs used were not directly relevant to the UK.

**Spolverato (2015)**<sup>104,105</sup>

Model structure	Multistate model with alternative model structures applied according to treatment received. The model considered the following states: undergoing liver resection or radiofrequency treatment (liver resection/RFA only), liver decompensation (liver resection/RFA only), HCC recurrence (liver resection/RFA only), progression of disease within Milan criteria (liver resection/RFA only), progressive disease outside Milan criteria, transplant waiting list, (post) liver transplant, and death.
Time horizon, perspective and discounting	Time horizon was not reported. Italian and US healthcare settings were considered using a payer perspective. Costs and benefits were discounted at a rate of 3%.
Population	Patients within the Milan criteria up to three tumours < 3 cm, or one tumour up to 5 cm.
Intervention and comparators	Liver transplant vs. liver resection or RFA with salvage liver transplantation.
Clinical evidence	Evidence was drawn from multiple sources identified in the literature and did not rely on comparative assessment of effectiveness. Outcomes modelled included: transplant wait time, post-transplant mortality, wait list dropout rate, liver decompensation, disease recurrence.
HRQoL	Health state utilities were based on data reported in Lim <i>et al.</i> <sup>109</sup> and did not vary by treatment received.
Resources and costs	Cost categories modelled included procedure costs, drug acquisition costs, disease management and patient follow-up costs. Resource data were drawn from two previous reviews of the literature: Cucchetti <i>et al.</i> <sup>99-101</sup> was used for Italian healthcare costs and Lim <i>et al.</i> <sup>109</sup> for US costs. Costs used were not directly relevant to the UK.

CLT, cadaveric liver transplantation; LDLT, living donor liver transplantation.

None of the identified studies considered a UK NHS perspective. One study considered an Italian setting only,<sup>99-101</sup> one a US setting only,<sup>102</sup> and one a Chinese setting only.<sup>103</sup> One further study considered both an Italian and a US setting.<sup>104,105</sup> In three studies the setting was not formally stated. In two of these studies,<sup>106-108</sup> costs were reported for a US setting, while a third study<sup>109</sup> reported costs for three alternative settings: USA, Singapore and Switzerland. All studies considered a payer perspective, where stated. As no study considered a UK setting, costs utilised are not relevant to the UK perspective. The identified studies are therefore unlikely to represent an informative source of resource data for any future economic evaluation adopting a UK perspective.

The model structures adopted in the identified studies varied significantly, with several alternative underlying approaches adopted. These included Markov models,<sup>103-106,109</sup> semi-Markov models,<sup>99-101</sup> and simulation approaches.<sup>107,108</sup> Model structures adopted were typically highly complex, with several using a large number of health states. Importantly, model structures did not conform to the three-state models commonly used in cancer evaluations. Despite a lack of consistency in the approach adopted across models, several features were common to the included studies. These included the modelling of recurrence of disease and the competing risks of declining liver function. Both of these features were uniquely associated with locoregional therapies such as RFA and resection and were not considered relevant to patients receiving a liver transplant. In several models, this meant that the structure adopted differed substantially between treatment arms.<sup>104-106,109</sup>

Because of the novel model structures adopted, treatment effects were often modelled using several parameters typically drawn from multiple studies. While this approach reflects the complex treatment pathways and allows a broader evidence base to be drawn upon, it comes with significant disadvantages. Namely, in this approach treatment effects are not based on comparative evidence and are highly likely to be subject to confounding biases. Further, while many models considered multiple outcomes, it is clear from model results that survival is the principal driver of benefits. An important consideration for future economic evaluations will therefore be how to best integrate available comparative evidence while also accounting for the divergent treatment pathways. In an ideal scenario this is likely to mean drawing directly on comparative evidence of survival. However, given the potentially curative nature of the evaluated treatments, such comparative evidence may be uninformative due to lack of maturity and developments in care for progressed HCC. It may therefore be necessary to draw on external data sources potentially linked to intermediate outcomes or events like transplant or recurrence of disease to populate an economic model.

## Model scope and availability of comparative data

Based on the systematic review of clinical effectiveness evidence and clinical advice, it is anticipated that there is a wide range of relevant comparators. These include established treatments such as resection and MWA, treatments that have more recently become available to UK patients such as SABR, and treatments that are no longer/rarely used in clinical practice (PEI and laser ablation). In principle, all of these therapies could be considered by a future cost-effectiveness analysis. However, clinical advice suggests that many of these newer technologies are rarely used in routine practice (e.g. ECT) owing to a lack of evidence/approval, while older technologies such as PEI and PAI have largely been discontinued due to lack of efficacy and concerns regarding AEs. Further, clinical advice suggests that some technologies such as IRE and SABR would not be used in the whole small-HCC population but instead would be reserved for patients with tumours in locations that are either difficult to treat or patients who are otherwise medically unsuitable for RFA. Any future economic analysis will therefore need to carefully consider the decision problem being addressed and which comparators are likely most relevant to decision-makers. Further, given the absence of evidence for some potentially relevant comparators, including many of the newer technologies, it may be necessary for a future economic analysis to focus only on a subset of all relevant comparators. This may limit the feasibility of implementing an informative economic analysis and is likely to impact on the strength of conclusions that can be drawn.

## Model structure and clinical data availability

The model structure typically adopted in economic evaluations of treatments for cancer uses a partition survival model (PSM) based around three health states: (1) pre progression, (2) post progression and (3) death. In a PSM the proportion of patients in each health state is determined directly from the survival curves, typically PFS and OS. Under this approach the proportion of patients in the 'pre-progression' state is determined by the PFS curve while the proportion in the 'post progression' state is determined by the difference between the modelled OS and PFS survival curves. Theoretically this approach could be adopted in the context of early HCC, but it may need adaptation to account for specific features of the indication. For example, as highlighted above, many models account for the competing risk associated with liver decompensation and potential for recurrence but not progressed disease. These complications may undermine the feasibility of a PSM approach, and the adaptations necessary may be easier to accommodate in a state transition model where it is often easier to explicitly acknowledge competing risks.

An alternative to the PSM approach would be to use a state transition model focused on utilising comparative evidence on recurrence and disease-free survival. This approach aligns with much of the previous cost-effectiveness literature and would more readily recognise the surrogate role that recurrence and disease-free survival play in determining OS. Under such an approach, post-recurrence survival would likely be modelled using a common set of assumptions for all treatments. While notionally this is a disadvantage as it assumes a consistent surrogate relationship between recurrence and OS, it would allow external data to be levied; this may provide improved estimates relative to the available trial data, which may be limited due to the short follow-up in many studies. This approach also allows post-recurrence survival to reflect recent developments in the treatment and care of patients with intermediate and advanced-stage HCC. This may be important given the more recent (post 2009) availability of sorafenib and other agents for the treatment of advanced HCC and the fact that the majority of the currently available clinical evidence is not from a UK setting.

Clinical advice on the aims of treatment emphasised the importance of recurrence, and particularly local recurrence, as a marker of treatment success. The importance of local recurrence as a determinant of mortality was also emphasised. It was, however, also emphasised that other factors are also important determinants of survival and may confound any relationship between local recurrence and OS. These included both intrahepatic and extrahepatic recurrence, which may lead to cancer progression regardless of local disease control. Further, clinicians noted the importance of liver function as a competing mortality risk, as well as its significance in determining patient quality of life.

This advice would appear to broadly support the use of a recurrence-focused approach but also emphasises the complexity of very early/early HCC and the need to account for the competing risks of disease progression and liver decompensation. The clinical data available to inform a recurrence-focused approach are, however, limited. Few studies identified in the clinical review reported recurrence, with only 10 of 27 identified studies reporting recurrence outcomes. This may impact on the feasibility of developing a robust economic model based around recurrence of disease, as it means that the totality of the evidence cannot be considered.

More broadly, inherent uncertainties in the clinical evidence, as well as concerns about the quality of included evidence, will have important consequences for any future economic analysis. As presented in *Updated network meta-analyses using RCT and non-RCT evidence*, current clinical evidence is insufficient to make recommendations about the relative effectiveness of the majority of treatments. An economic analysis cannot resolve these uncertainties and will necessarily be limited by them. Importantly, these uncertainties are likely to undermine the ability of any future economic analysis to make recommendations about which treatments are most cost-effective. This may undermine the value of implementing an economic analysis. An economic analysis may, however, still be worthwhile because of its ability to quantify the uncertainty associated with implementation decisions. In doing so, an economic analysis can help provide information about the value of future research; see [Value of information](#) below for further discussion.

### **Utilities and quality of life**

In the literature identified as part of the clinical effectiveness review, no RCTs and only one non-RCT collected quality-of-life data,<sup>87</sup> and no study reported utility data. Any new economic evaluation will therefore have to identify alternative sources of relevant utility data. The identification of relevant utility data is likely to require detailed searches of the literature. Based on the cost-effectiveness evidence identified in our review, several studies reported utility values that may be relevant to any future analysis.<sup>99-103,106,109</sup> However, the provenance of some of the values reported is unclear.<sup>109</sup> In other cases, it is also apparent that the values obtained are taken from patients with liver disease rather than specifically from patients with HCC.<sup>102,103</sup> Further, several of the evaluations identified in the cost-effectiveness review highlighted limitations in the available quality-of-life data.<sup>102,104,105,109</sup> Identifying relevant utility data is likely to represent a significant challenge and source of uncertainty for any new economic evaluation in early HCC.

### **Resource use and costs**

Resource use and costs should include treatment costs (acquisition, procedures, and monitoring), changes in health service utilisation driven by disease status (i.e. progression-free, progressed disease, and death), and AE management. Costing data from previous economic analyses in early HCC are unlikely to be informative due to differences in perspective; no study was conducted from a UK perspective. Further, few studies reported relevant resource-use estimates associated with specific treatments. Previous economic evaluations are therefore unlikely to provide resource inputs for a new model.

Several of the studies identified in the clinical effectiveness review reported on useful economic outcomes such as length of hospital stay. Assuming these studies are generalisable to a UK setting, these outcomes could be used to support inputs regarding acute care and monitoring following treatment. The majority of resource-use inputs will, however, need to be identified in further research. This may be in the form of a clinician survey to elicit resource utilisation or identification of relevant costing studies. Alternatively, health state management costs may be informed by previous UK economic evaluations in advanced HCC, and adapted to account for the target early HCC population. Costing data for the UK are readily available from several commonly used sources. These include NHS reference costs,<sup>112</sup> Personal Social Services Research Unit,<sup>113</sup> and the British National Formulary.<sup>114</sup> While further research is necessary, the availability of resource-use and costing data is unlikely to represent a significant barrier to implementing a future economic evaluation.

### **Value of information**

The construction of a *de novo* economic analysis in which uncertainty is fully parameterised would allow the implementation of a VOI analysis. A VOI analysis permits the value of reducing decision uncertainty to be quantified in monetary terms. The VOI can then be compared with the costs of further studies and used to assess whether additional research should be conducted to reduce decision uncertainty.

In the context of the current evidence, a VOI analysis may be particularly helpful, as there are currently several treatment alternatives for which there is limited evidence on effectiveness. A VOI analysis could help prioritise which of these treatments should be assessed in future trials, accounting for both the degree of clinical uncertainty and the economic case for a specific treatment. This may be of particular relevance in considering treatments that are currently rarely used in NHS practice but may be effective; for example, laser ablation and RFA. Moreover, a VOI analysis may help to provide clearer guidance on where research is not worthwhile despite the presence of clinical uncertainty. For example, VOI may be able to rule out particularly expensive technologies on cost grounds alone despite the potential for clinical benefit.

## Chapter 6 Patient and public involvement

### Aim

The aim of patient and public involvement was to ensure that the patient's perspective was captured at all stages, from protocol development through to interpreting the results of the project and drawing conclusions and recommendations for further research.

### Methods

A patient collaborator was recruited to the project at the proposal writing stage via 'Involvement@York', the patient and public involvement network at the University of York. The patient collaborator attended all advisory group meetings and provided ongoing advice throughout the project. The patient collaborator was also consulted when producing materials in 'plain English', such as materials used when recruiting additional patients to the advisory group and the plain English summary section of the final report. The patient collaborator will be consulted during further dissemination activities.

Four additional patients were identified by our clinical advisors and recruited as members of the advisory group. With help from the patient collaborator, a lay summary of the project was produced describing the project, the role of advisory group members and details of how patients would be compensated for their time. This was circulated to patients who had expressed an interest in being a member of the advisory group. Patients were also provided with a lay summary of the different interventions included in the systematic review.

One member of the project team (RW) was the main contact for all patient advisors and held individual meetings with patients at the protocol development stage. During this initial meeting, patients were given background information to the project and a rudimentary description of the protocol and were asked for their comments, specifically whether any patient-relevant outcomes or aspects of treatment were missing from the protocol. Owing to the COVID-19 pandemic, all advisory group meetings were held via the Zoom™ online videoconferencing platform (Zoom Video Communications, San Jose, CA, USA), rather than in person. Patients were invited to attend the next advisory group meeting and the end-of-project workshop (see *Workshop*). Patients were also asked to comment on the final report.

### Results

All four patients were available at the beginning of the project to advise on the protocol. The patient collaborator and three of the patient advisory group members attended the second advisory group meeting held midway through the project to discuss the interim findings, prioritise interventions for further review and prioritise the most relevant patient outcomes. Patients provided helpful information about the outcomes most important to them, such as length of hospital stay and disruption to life (interventions requiring multiple appointments or repeat treatments) and level of pain involved. Non-recurrence of disease was another important outcome to patients. The patient collaborator and two patient advisory group members attended one of the end-of-project workshops. Unfortunately the other two patients were unavailable around the time of the workshops; in view of the reasons for their lack of availability, they were not pursued to attend at a different time. Patients were surprised by the lack of data on patient preference and quality-of-life outcomes in the existing evidence base. The patient collaborator and two patients commented on the final report.

## **Discussion and conclusions**

The patient and public involvement aspect of the project highlighted the outcomes most important to patients, which informed the development of the data extraction form. Their views added context to the review findings and their input was valuable when drawing conclusions and making recommendations for further research. The initial meeting with patients was informative to help the researchers understand the experience of patients, their concerns and preferences.

## **Reflective/critical perspective**

Patient involvement was a valuable part of this project, enabling researchers to understand important aspects of the different treatment options from a patient's perspective. One drawback was that meetings had to be held via the Zoom online videoconferencing platform, owing to the COVID-19 pandemic, which constrained the interactions with patients.

The feedback from patients was positive; they commented that information was presented clearly and that they found the meetings interesting and enjoyed being involved in the project.

## Chapter 7 Workshop

Two workshops were held with clinical and patient advisory group members and additional clinicians with an interest in HCC (identified by advisory group members) in order to discuss the project findings and identify key priorities for future research. Due to the COVID-19 pandemic, the workshops were held using the Zoom online videoconferencing platform, on 29 November and 2 December 2021. Prior to the workshops, the attendees were sent a short summary of the findings of the project to date, including a summary of the methods and results of the systematic reviews, NMAs and the assessment of the feasibility of economic modelling.

Members of the project team presented a summary of the findings of the project and responded to clarification questions. There was a general discussion of the interpretation of the project findings, and workshop participants were asked about the key priorities for future research, including interventions, patient groups, and outcomes.

The lack of evidence for many interventions, low quality of the available evidence and uncertainty of the findings was highlighted. The generalisability of the findings of studies from East Asia (where the underlying aetiology of the liver disease differs from that in the West) was discussed, since most of the RCTs assessed RFA, which is more widely used in the East, whereas MWA has become the standard of care in most centres in the West. It was agreed that differences in underlying liver disease are likely to affect the absolute OS of patients, rather than the relative survival when comparing one treatment against another.

The progression in the West from RFA to MWA as the standard of care has been driven by technological advances and ease of use of MWA (which only requires single needle placement, so is both faster and simpler to deliver) rather than data on improved clinical effectiveness. MWA gives a more predictable ablation zone up to 3 cm, whereas the RFA ablation zone is less predictable towards the periphery. However, it was considered that, moving forward, it would not be appropriate to compare the clinical effectiveness of RFA versus MWA, as many interventional radiologists in the UK only know how to use MWA, not RFA, and clinicians believe that MWA is the superior treatment, so it may be difficult to recruit patients to a trial comparing the two treatments. In addition, RFA is only used for tumours up to 2 cm (owing to increased local recurrence after RFA in lesions larger than 2 cm), whereas MWA can be used for larger tumours; therefore, any trial comparing both technologies would have to restrict recruitment to patients with tumours up to 2 cm in order for patients to be eligible for both treatment arms. Lesions close to hepatic vessels are also less amenable to RFA, reducing the eligible patient cohort further.

RFA was used as the baseline treatment in the NMA (for comparison against other treatments) because it was the most widely assessed technology in the RCTs. Historically, surgery was considered to be the gold standard, before RFA became available. However, it would not be appropriate to compare the effectiveness of surgery versus ablation, as the risks of surgery for many patients are too high. Resection is not suitable for cirrhotic patients with marginal liver function or patients with clinically significant portal hypertension. Tumour location is also important; resection would not be suitable for patients with central tumours, particularly in patients with cirrhosis, as the risks of resection are much higher than those of ablation. At the second workshop, the comparison of MWA versus resection was discussed further, as some centres are still quite 'surgery heavy' and may want to see more trial evidence on MWA versus resection, although most centres are moving towards MWA owing to the complications of resection making it less acceptable than ablation.

Specific effectiveness outcomes and the association between liver decompensation (liver failure) and mortality were discussed. Registry data suggest that around half of patients who undergo ablation will die following liver decompensation, and half will die without liver decompensation; therefore, the risk

between recurrence and mortality is important, but there is substantial competing risk according to the severity of underlying disease. At the second workshop it was highlighted that some HCC treatments have a risk of causing liver decompensation, although patients with early HCC have good Child–Pugh scoring and good performance status.

It is difficult to demonstrate treatment benefit in a trial when there are competing risks of both liver disease progression to decompensation in addition to the risks of recurrent or new HCC. Recurrence can be local (near the site of the previously ablated lesion) or distant (a new lesion elsewhere within the liver); therefore, treatment of the original tumour may not impact on OS. Rates of recurrent and new HCC are very high; up to half of patients will have a new metachronous cancer within 3 years of treatment of the index lesion, which is a further driver of poor outcomes in this patient group. Rates of metachronous disease would be expected to be similar between different treatment arms in a trial, unless one of the interventions treats the whole liver. This is why transplantation is theoretically the best treatment for early-stage HCC, because it replaces the liver that has malignant potential with a new one, so there is no longer the risk of metachronous disease or decompensation. However, liver transplant is not normally the primary intervention for the population of patients with early-stage HCC.

Quality of life was only reported in one included study, a non-randomised study undertaken in China. It is important to assess quality-of-life outcomes and patient acceptability in any future trial; there is a lack of evidence on these important outcomes in the existing literature.

The problem of patient recruitment was discussed, as there are not many patients with early-stage ( $\leq 3$  cm) HCC in the UK. The marginal benefit of novel treatments compared with the existing standard of care is likely to be small, so future studies would need to be large to demonstrate a significant difference in outcomes. Therefore, an international multicentre RCT may be more appropriate than a UK-based trial.

At the first workshop, clinicians said that SABR and proton beam therapy are interventions of interest and that a trial of SABR or proton beam therapy versus MWA would be useful, although use of proton beam therapy is limited by geographical availability. Local control rates with both treatments are very high; therefore, undertaking a trial that was sufficiently powered to show a survival benefit would be difficult, since neither deals directly with recurrence (metachronous or extrahepatic disease) and neither has an impact on rates of decompensation. Local recurrence would have to be the primary outcome in such a trial, with OS and PFS as secondary outcomes. At the second workshop it was also agreed that local recurrence and overall recurrence are both important outcomes.

It is internationally recognised that there needs to be more trial-based evidence for SABR in the treatment of patients with early-stage HCC. The availability of such evidence is limited by the fact that ablation techniques such as MWA and RFA are usually employed first for patients with early HCC and SABR reserved for recurrent, refractory or more advanced disease. SABR can usually only be delivered once because of the radiation dose, whereas ablation can be repeated; therefore, there is also the question of when it should be used – should it be saved until later in the treatment pathway? There was discussion around assessing different treatment sequences. Although treatment sequencing is an important question, the difficulty is the heterogeneity of recurrence, which has implications for the next treatment choice; therefore, it may not be possible to predetermine the second line in the sequence. Both MWA and SABR can be used for patients with tumours  $\leq 3$  cm, and both interventions can be used to treat more than one lesion at once; therefore, trial eligibility criteria would have to reflect this. There would also need to be eligibility criteria limitations based on liver function, as patients with advanced/moderately advanced liver disease are not suitable for SABR but could possibly be suitable for ablation; patients recruited to a trial would have to be eligible for both treatments.

At the second workshop, clinicians considered that a trial of SABR versus MWA may be less appropriate; MWA is a good treatment for small tumours, while SABR is usually reserved for tumours that do

not respond or are unsuitable for ablation or are in a difficult location. Therefore, SABR and other radiotherapy techniques would not replace MWA. In addition, MWA can be repeated, while SABR can usually only be used once. This positioning of SABR reflects current NHS England commissioning guidance which suggests SABR and MWA for different patient populations; for patients with very early/early-stage HCC, ablation is the first choice, while if the lesion cannot be clearly visualised for ablation or is in a location that cannot be reached with a needle, then TACE would be offered. If TACE is contraindicated (e.g. for cardiac reasons or if the patient has had TACE previously and failed), then SABR would be offered. The clinicians also noted there is a study in North America comparing SABR versus proton beam therapy; proton beam therapy may be the preferred modality for patients depending on disease location.

At the first workshop, clinicians said that IRE can be used for lesions that are very central; therefore, a trial of IRE versus MWA for the subgroup of patients with central lesions may be useful, although SABR could also be used. At the second workshop it was agreed that IRE is sometimes used for more challenging tumours, but owing to the evidence base being very limited for IRE, SABR is the preferred option. In addition, IRE is quite costly; therefore, MWA would be used when suitable; they would not be comparable in a trial.

ECT is very similar to IRE but with the addition of bleomycin. It is beginning to feature in Europe, so may be of interest.

Cryoablation was not considered to be of interest as it is a high-risk treatment. It has not been widely adopted in the West. At the second workshop, it was stated that more evidence is being published on cryoablation, especially for lesions that are difficult to treat with MWA, such as those that are near the dome of the liver or close to the heart, where freezing therapy is slightly less damaging than heat treatment; thus it is mostly used for those lesions that are difficult to treat because of nearby vital structures. However, if IRE is available, that would be used rather than cryoablation, so cryoablation has a lower priority.

Laser has also not been widely adopted in the West. It involves multiple needle placement, whereas MWA only requires single needle placement so is both faster and simpler to deliver. However, there are no clinical effectiveness data comparing it with MWA, so the comparative effectiveness is unknown. However, ease of use is an important consideration in treatment choice; any intervention with a substantial learning curve barrier is going to be less easily accepted from a clinical perspective. HIFU has also been around for several years but has not been widely adopted.

Histotripsy is currently being evaluated as an investigational product; therefore, it should not be assessed further until efficacy has been demonstrated. However, it appears to be very promising and may be of interest further down the line.

At the second workshop, it was considered that the questions to answer in early HCC are more in the setting of challenging locations, less fit patients and in the setting of incomplete response to primary therapy, rather than a comparison with the current preferred first treatment option. There is probably some variation between multidisciplinary teams on whether they would offer TACE and whether they have SABR and/or IRE available. It may be difficult to define the population and ensure that a trial was acceptable to multidisciplinary teams that might have slight variations in practice and also looking at what technologies are available locally to a patient.



# Chapter 8 Discussion

## Summary of findings

The aim of this research was to evaluate and compare the effectiveness of ablative and non-surgical therapies for patients with small (up to 3 cm) HCC tumours. The key objectives were: to systematically identify all RCTs of ablative and non-surgical therapies for HCC; to evaluate their quality and applicability to UK populations; to determine the comparative effectiveness of therapies using NMA techniques; to supplement the RCT evidence with non-randomised prospective comparative studies of specific therapies where the evidence base was insufficient; to identify priority areas where additional high-quality evidence is required; and to assess whether future economic analysis based on the findings would be feasible and worthwhile.

Thirty-seven RCTs (one ongoing, 36 completed) were included in the systematic review. Several included patients with tumours larger than 3 cm, but reported separate results for the subgroup of patients with tumours up to 3 cm, although often the data reported for the subgroup were limited to response and/or AE outcomes. The RCT evidence was limited; most studies were small and at a high RoB (12 RCTs) or had some bias concerns (14 RCTs). The vast majority of RCTs were conducted in China or Japan, which has implications for the generalisability of results to the UK population, owing to differences in HCC aetiology and the different treatment options for the underlying liver disease. The most frequently assessed ablative therapy was RFA, which is widely used in Asia. However, in the UK and Europe MWA has been more widely adopted because of advances in microwave technology; MWA gives a more predictable ablation zone and is easier and faster to use, requiring single needle placement. Many interventional radiologists in the UK do not have experience of using RFA.

The results of many of the included RCTs were heterogeneous, particularly for the comparison of RFA versus surgical resection, with some RCTs favouring surgical resection and others favouring RFA or reporting similar OS and disease-free survival rates between treatment groups. However, AE rates were higher after resection. There was no evidence to suggest a difference between treatment with RFA and resection in the NMA.

Data comparing RFA with MWA, laser ablation or proton beam therapy were limited, with few RCTs and very small sample sizes. RCTs assessing RFA in combination with other treatments were also limited by small sample sizes. The uncertainty associated with the available data is demonstrated in the NMA results, where CIs were generally wide and most crossed the line of no effect. The estimated treatment effectiveness ranking was also very uncertain, with very wide CIs for most interventions.

The only firm conclusion that can be drawn from the available RCT data is that RFA appears to be better than PEI in terms of OS, PFS and recurrence. However, AEs appear to be more frequent after RFA than PEI, although this outcome could not be evaluated in a NMA. PAI appears to have similar effectiveness to PEI and had marginally worse PFS than RFA in the NMA, although data for this comparison were more limited.

One trial assessed RFA in combination with iodine-125, which appeared to be superior to RFA in terms of OS and overall recurrence; however, clinical advisors stated that this is only used in selected centres in China, and very few centres outside of China have used this combination.

No RCT evidence was identified for several of the interventions of interest: HIFU, cryoablation, IRE, ECT, histotripsy, SABR and wider radiotherapy techniques. As highlighted at the project workshop, histotripsy is currently being evaluated as an investigational product; therefore, it is unlikely that randomised evidence will be available within the next few years. Cryoablation, IRE, ECT and SABR are generally

reserved for the subgroup of patients with lesions that are more challenging to treat because of their location. SABR is also reserved in the treatment pathway for patients with recurrent, refractory or more advanced disease, or patients with comorbidities that make them unsuitable for ablative therapies. This makes it more difficult to undertake a randomised trial, as recruited patients would have to be eligible for both treatment arms. In addition, some of these technologies are less widely available and have a higher cost than ablative technologies such as RFA and MWA.

The threshold analysis suggested that additional evidence could plausibly change the NMA result for comparisons including RFA, MWA or laser ablation, as well as RFA in combination with TACE, systemic chemotherapy or iodine-125. Therefore, a systematic review of non-randomised prospective comparative studies was undertaken to identify evidence on RFA, MWA, laser ablation, HIFU, cryoablation, IRE, ECT, histotripsy, SABR and wider radiotherapy techniques, compared with each other or with surgical resection.

The systematic review of non-randomised evidence included 14 studies, although only two studies did not have a high RoB. Several studies allocated patients to treatment groups based on tumour characteristics (such as tumour location), meaning that there were differences in baseline characteristics between treatment groups that could be prognostic. This has implications for the interpretation of the non-randomised evidence; in addition, included patients may not have been eligible for both of the treatments assessed. Again, the vast majority of studies were conducted in China or Japan, with implications for the generalisability of results to the UK HCC patient population. In view of the significant limitations of the non-randomised studies, the studies with a high RoB were not included in the updated NMAs, leaving only the two studies that had a low RoB or some bias concerns. Additional non-randomised comparative data from Leeds Teaching Hospitals NHS Trust were made available by one of the clinical advisors, prior to publication; these data were also included in the updated NMAs.

The results of the updated NMAs including the non-randomised evidence were largely consistent with those of the NMAs of RCTs. As with the NMAs of randomised evidence, the findings were highly uncertain. However, the results suggested that MWA appears to be better than PEI and PAI in terms of OS and PFS. Resection appears to be better than PEI in terms of OS, and better than PEI and PAI in terms of PFS. In addition, IRE appears to be worse than RFA and RFA + PEI in terms of local recurrence.

The feasibility of developing an economic analysis to inform decision-makers on the cost-effectiveness of alternative treatments for small HCCs was assessed. This included a targeted literature review, which was undertaken to identify previous economic evaluations in very early/early HCC. The key features of the identified studies were summarised and used to inform the development of a conceptual model and to consider the data needed to develop a robust economic analysis. The review identified that previous economic evaluations have used recurrence events and liver function to predict long-term outcomes. This approach is likely to be the most appropriate way to model early HCC given the current evidence. Limitations in the available clinical data are, however, likely to impact on the feasibility of developing a robust economic analysis and limit any conclusions that could be drawn. Specifically, uncertainties in the clinical effectiveness will pervade any future economic analysis.

Given these uncertainties, a VOI analysis may be helpful and could help prioritise which of these treatments should be assessed in future trials, accounting for both the degree of clinical uncertainty and the economic case for a specific treatment. This may be of particular relevance in considering effective treatments that are currently rarely used in NHS practice.

There are considerable limitations to the existing evidence base on ablative and non-surgical therapies for early HCC. Two workshops were held to discuss the project findings and identify key priorities for future research; three patients and six clinicians provided expert advice. In view of the wide adoption of MWA as the standard of care within the UK and Europe, it was agreed that MWA would be the most appropriate comparator in any future trials. Clinicians considered that ablative technologies that are

more complex and take longer to deliver than MWA (e.g. laser and RFA, which require multiple needle placement) are unlikely to displace MWA as the preferred ablative therapy, despite a lack of clinical effectiveness evidence demonstrating better outcomes.

Specific interventions considered to be of particular interest to the HCC community were SABR and proton beam therapy, although these radiotherapy-based treatments can usually only be delivered once, whereas ablation can be repeated. In addition, there would need to be limitations to trial eligibility criteria, as patients with advanced/moderately advanced liver disease are not suitable for SABR because of the radiotherapy dose delivered to the surrounding liver. SABR and other radiotherapy techniques are unlikely to replace MWA as the first treatment choice; these techniques are generally reserved for a subgroup of patients depending on their suitability for ablation, tumour location and other patient and disease characteristics. A trial of IRE versus MWA for the subgroup of patients with central lesions may be useful, although again IRE would be unlikely to replace MWA in patients suitable for ablation. A trial of ECT versus MWA was also considered to be of interest. For early HCC, further research may be most relevant in the setting of challenging locations, less fit patients and incomplete response to primary therapy, rather than a comparison with the current preferred first treatment option (MWA).

Histotripsy was identified as an investigational product that may be promising in the future; however, it is at an early stage of regulatory approval, so should not be assessed until efficacy has been demonstrated.

Because of the low number of patients in the UK with early-stage HCC who would be eligible for all treatments within a trial, particularly for those interventions reserved for the subgroup of patients with more challenging tumours, it is likely to be more feasible to undertake an international multicentre RCT than a UK-based trial, in terms of recruiting sufficient patients to demonstrate a significant difference in outcomes. However, patients' disease characteristics, such as aetiology of liver disease and prior treatments received, would need to be similar to those of HCC patients in the UK to ensure that trial results were generalisable to the UK HCC population. Unfortunately, there were insufficient data on specific patient subgroups (i.e. tumour size and number, severity of cirrhosis and underlying liver disease) to enable subgroup analysis to be undertaken within the review. Therefore, it is unclear whether these characteristics are effect modifiers.

Local recurrence, overall recurrence, OS, PFS and HRQoL are important outcomes that should be assessed in any future trials. The definition of specific outcomes, such as recurrence and PFS, should be consistent in future trials to allow results to be compared and synthesised in the future.

The 2022 update of the BCLC strategy for prognosis prediction and treatment recommendation states that further prospective studies are needed to define the role of SABR for very early HCC.<sup>7</sup>

## Strengths and limitations

The key strengths of this assessment are the comprehensive searches for relevant RCT evidence, the systematic data extraction and assessment of the quality and applicability of the included studies, and the inclusion of relevant data in NMAs of four important clinical effectiveness outcomes in an attempt to draw indirect comparisons of the therapies and rank them from best to worst in terms of the relevant outcomes.

The systematic review of RCTs was supplemented with a targeted review of non-randomised evidence in an attempt to fill gaps in the RCT evidence base and strengthen the evidence where data on specific comparisons were considered to be weak. Attention was focused on those interventions with current clinical relevance and those comparisons sensitive to potential changes in the evidence, as determined using novel threshold analysis techniques.

## DISCUSSION

The project benefited from the expertise of several patient and clinical advisors, with meetings held at key stages of the project. However, the absence of 'in person' meetings, owing to the COVID-19 pandemic, constrained the interactions with patients. In addition, two of the patient advisory group members were unfortunately unavailable for the workshops at the end of the project.

The assessment was limited by the weaknesses in the clinical evidence base. There was no evidence on several of the interventions of interest, and the evidence was extremely weak (in terms of size and quality) for most of the other therapies, limiting our ability to draw any firm conclusions. Because of the significant gaps in the evidence base, the recommendations for prioritising specific therapies and comparisons for future research were primarily made based on expert advice received during the end-of-project workshop.

# Chapter 9 Conclusions

## Implications for practice

The evidence on ablative and non-surgical therapies for early and very early HCC is very limited. The only firm conclusions that can be drawn from the available data are that PEI and PAI are inferior to RFA, and that they also appear to be inferior to MWA and resection, for certain survival outcomes. There is insufficient evidence to draw any conclusions on quality-of-life outcomes.

The uptake of specific ablative therapies in the UK appears to be based more on technological advancements and ease/speed of use (and NHS England commissioning policies) than on high-quality evidence demonstrating superior clinical effectiveness of one therapy over another.

## Recommendations for research

There are currently no comparative data on several ablative and non-surgical therapies, particularly those treatments reserved for the subgroup of patients with more challenging tumours. However, owing to the small number of such patients who would be eligible for both treatment arms within a trial, it is likely to be difficult to recruit sufficient numbers of patients to demonstrate a significant survival benefit, particularly in the presence of a competing risk of recurrence from the underlying liver disease.

Future studies should assess local recurrence, overall recurrence, OS, PFS, HRQoL and patient acceptability, using clear and consistent definitions, in order to allow results to be compared across studies.

It is difficult to make firm recommendations for research based on our findings. The current evidence suggests a trial of MWA versus RFA versus resection could address uncertainty about the standard of care; however, clinicians consider this unlikely to be helpful as RFA is no longer widely used in NHS practice.

Clinical experts suggest that SABR is a promising intervention and could be compared with MWA; this may have international relevance, allowing for wider patient recruitment through multinational trials. However, SABR can usually only be used once because it is limited by the radiotherapy dose received by the surrounding liver, so further research is needed to identify where it should sit in the treatment pathway.

There were insufficient data on specific patient subgroups (i.e. relating to tumour size and number, severity of cirrhosis and underlying liver disease) to enable subgroup analysis to be undertaken. Therefore, further research to assess whether certain disease characteristics may modify treatment effect could be beneficial.

Feasibility studies could address these potential issues and complexities in undertaking research in this area prior to undertaking a trial. This would enable investigation of: the acceptability of the intervention (and comparator) to both clinicians and patients and their willingness to participate in a trial; the practicality of delivering the intervention; and the ability to measure relevant outcomes.



# Acknowledgements

We would like to thank the members of our advisory group: Dr Rebecca Goody, Consultant Clinical Oncologist, Leeds Teaching Hospitals NHS Trust; Dr Jai Patel, Consultant Vascular and Interventional Radiologist, Leeds Teaching Hospitals NHS Trust; Professor Ajith Siriwardena, Consultant Hepatobiliary and General Surgeon, Manchester University NHS Foundation Trust; Dr Tze Wah, Consultant Interventional Radiologist, Leeds Teaching Hospitals NHS Trust; and the patient advisors: Dave Clarke, Ian Doyle, Richard McCabe and Ian Teunion. In addition, we would like to thank the other clinicians who attended the workshop: Dr Salil Karkhanis, Consultant Interventional Radiologist, University Hospitals Birmingham NHS Foundation Trust; and Professor Derek Manas, Professor of Hepatobiliary and Transplant Surgery, Newcastle Hospitals NHS Foundation Trust. We would also like to thank Dr Tze Wah for providing additional unpublished registry data that were included in the updated network meta-analyses. We are grateful for the input of CRD staff early in the life of the project who contributed to the proposal or protocol, namely Lindsay Claxton and Kath Wright. We would also like to thank Jinshuo Li and Nick Meader for assistance translating Chinese and Japanese publications.

## Contributions of authors

**Ros Wade** (<https://orcid.org/0000-0002-8666-8110>) (Research Fellow) contributed to the protocol, study selection, data extraction, validity assessment and synthesis of the included studies. She also contributed to the interpretation of the results and the writing of the report.

**Emily South** (<https://orcid.org/0000-0003-2187-4762>) (Research Fellow) contributed to the selection, data extraction, validity assessment and synthesis of the included studies. She also contributed to the interpretation of the results and the writing of the report.

**Sumayya Anwer** (<https://orcid.org/0000-0002-1740-0399>) (Research Fellow) undertook the threshold analysis and incorporated non-randomised evidence into the network meta-analyses. She also contributed to the interpretation of the results and the writing of the report.

**Sahar Sharif-Hurst** (<https://orcid.org/0000-0001-6885-0456>) (Research Fellow) contributed to the protocol, study selection, data extraction and validity assessment. She developed the synthesis models and undertook the analyses. She also contributed to the interpretation of the results and the writing of the report.

**Melissa Harden** (<https://orcid.org/0000-0003-2338-6869>) (Information Specialist) contributed to the protocol, developed search strategies, conducted a range of searches to locate studies and wrote the sections of the report relating to the literature searches.

**Helen Fulbright** (<https://orcid.org/0000-0002-1073-1099>) (Information Specialist) developed search strategies, conducted a range of searches to locate studies and contributed to the sections of the report relating to the literature searches.

**Robert Hodgson** (<https://orcid.org/0000-0001-6962-2893>) (Research Fellow in Health Economics) undertook the review of economic studies. He was responsible for assessing the feasibility of economic modelling and contributed to the writing of the report.

**Sofia Dias** (<https://orcid.org/0000-0002-2172-0221>) (Professor in Health Technology Assessment) provided methodological expertise throughout the project and contributed to the protocol, interpretation of the results and the writing of the report.

## ACKNOWLEDGEMENTS

**Mark Simmonds** (<https://orcid.org/0000-0002-1999-8515>) (Senior Research Fellow) provided methodological expertise throughout the project and contributed to the protocol, interpretation of the results and the writing of the report.

**Ian Rowe** (<https://orcid.org/0000-0003-1288-0749>) (Honorary Consultant Hepatologist) provided expert clinical advice, contributed to the protocol and interpretation of the results, and commented on drafts of the report.

**Patricia Thornton** (<https://orcid.org/0000-0002-8814-0790>) (patient collaborator) provided advice as a patient expert, contributed to the protocol and interpretation of the results, and commented on drafts of the report.

**Alison Eastwood** (<https://orcid.org/0000-0003-1079-7781>) (Professor of Research) contributed to the protocol, study selection and synthesis of the included studies. She also contributed to the interpretation of the results and the writing of the report. Alison had overall responsibility for the project.

## Publications

South E, Wade R, Anwer S, Sharif S, Harden M, Fulbright HA, *et al.* The effectiveness of ablative and non-surgical therapies for early hepatocellular carcinoma: systematic review and network meta-analysis of randomised controlled trials. *Cancer Medicine* 2023. <https://doi.org/10.1002/cam4.6643>

## Data-sharing statement

The majority of data are available in the appendices to the report; any additional information can be obtained from the corresponding author.

## Ethics statement

This systematic review did not require ethical approval.

## References

1. Cancer Research UK. *Liver cancer statistics*. URL: <https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/liver-cancer> (accessed 7 December 2021).
2. European Association for the Study of the Liver. EASL clinical practice guidelines: management of hepatocellular carcinoma. *J Hepatol* 2018;**69**:182–236.
3. Burton A, Tataru D, Driver RJ, Bird TG, Huws D, Wallace D, *et al*. Primary liver cancer in the UK: incidence, incidence-based mortality, and survival by subtype, sex, and nation. *JHEP Rep* 2021;**3**:100232. <https://doi.org/10.1016/j.jhepr.2021.100232>
4. Global Burden of Disease Liver Cancer Collaboration. *The burden of primary liver cancer and underlying etiologies from 1990 to 2015 at the global, regional, and national level: results from the Global Burden of Disease study 2015*. *JAMA Oncol* 2017;**3**:1683–91. <https://doi.org/10.1001/jamaoncol.2017.3055>
5. Balogh J, Victor D, Asham EH, Burroughs SG, Boktour M, Saharia A, *et al*. Hepatocellular carcinoma: a review. *J Hepatocell Carcinoma* 2016;**3**:41–53. <https://doi.org/10.2147/jhc.S61146>
6. National Institute for Health and Care Excellence. *Cirrhosis in Over 16s: Assessment and Management [NG50]*. London: NICE; 2016.
7. Reig M, Forner A, Rimola J, Ferrer-Fàbrega J, Burrel M, Garcia-Criado A, *et al*. BCLC strategy for prognosis prediction and treatment recommendation: the 2022 update. *J Hepatol* 2021;**76**:681–93. <https://doi.org/10.1016/j.jhep.2021.11.018>
8. Yang JD, Heimbach JK. New advances in the diagnosis and management of hepatocellular carcinoma. *BMJ* 2020;**371**:m3544. <https://doi.org/10.1136/bmj.m3544>
9. NHS England. *Clinical Commissioning Policy Stereotactic Ablative Radiotherapy (SABR) for Hepatocellular Carcinoma (Adults) [URN: 1913]*: Leeds: NHS England; 2019.
10. National Institute for Health and Care Excellence. *Irreversible Electroporation for Primary Liver Cancer. Interventional Procedures Guidance [IPG664]*. London: NICE; 2019.
11. National Institute for Health and Care Excellence. *Radiofrequency Ablation of Hepatocellular Carcinoma. Interventional Procedures Guidance [IPG2]*. London: NICE; 2003.
12. National Institute for Health and Care Excellence. *Microwave Ablation of Hepatocellular Carcinoma. Interventional Procedures Guidance [IPG214]*. London: NICE; 2007.
13. Awad T, Thorlund K, Gluud C. Cryotherapy for hepatocellular carcinoma. *Cochrane Database Syst Rev* 2009;**4**:CD007611. <http://dx.doi.org/10.1002/14651858.CD007611.pub2>
14. Majumdar A, Roccarina D, Thorburn D, Davidson BR, Tsochatzis E, Gurusamy KS. Management of people with early- or very early-stage hepatocellular carcinoma. *Cochrane Database Syst Rev* 2017;**3**:CD011650. <http://dx.doi.org/10.1002/14651858.CD011650.pub2>
15. Weis S, Franke A, Mössner J, Jakobsen JC, Schoppmeyer K. Radiofrequency (thermal) ablation versus no intervention or other interventions for hepatocellular carcinoma. *Cochrane Database Syst Rev* 2013;**12**:CD003046. <http://dx.doi.org/10.1002/14651858.CD003046.pub3>
16. Weis S, Franke A, Berg T, Mössner J, Fleig WE, Schoppmeyer K. Percutaneous ethanol injection or percutaneous acetic acid injection for early hepatocellular carcinoma. *Cochrane Database Syst Rev* 2015;**1**:CD006745. <http://dx.doi.org/10.1002/14651858.CD006745.pub3>
17. Kamarajah SK, Bundred JR, Littler P, Reeves H, Manas DM, White SA. Treatment strategies for early stage hepatocellular carcinoma: a systematic review and network meta-analysis of randomised clinical trials. *HPB (Oxford)* 2020;**11**:1–11.

18. Tian G, Yang S, Yuan J, Threapleton D, Zhao Q, Chen F, *et al.* Comparative efficacy of treatment strategies for hepatocellular carcinoma: systematic review and network meta-analysis. *BMJ Open* 2018;**8**:e021269. <https://doi.org/10.1136/bmjopen-2017-021269>
19. Lan T, Chang L, Rahmathullah MN, Wu L, Yuan YF. Comparative efficacy of interventional therapies for early-stage hepatocellular carcinoma: a PRISMA-compliant systematic review and network meta-analysis. *Medicine (Baltim)* 2016;**95**:e3185. <https://dx.doi.org/10.1097/MD.0000000000003185>
20. Park JW, Chen M, Colombo M, Roberts LR, Schwartz M, Chen PJ, *et al.* Global patterns of hepatocellular carcinoma management from diagnosis to death: the BRIDGE Study. *Liver Int* 2015;**35**:2155–66. <https://doi.org/10.1111/liv.12818>
21. Centre for Reviews and Dissemination. *Systematic Reviews: CRD's Guidance for Undertaking Reviews in Health Care*. York: University of York; 2009.
22. Moher D, Liberati A, Tetzlaff J, Altman DG; The PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ* 2009;**339**:b2535. <https://doi.org/10.1136/bmj.b2535>
23. Hutton B, Catalá-López F, Moher D. The PRISMA statement extension for systematic reviews incorporating network meta-analysis: PRISMA-NMA. *Med Clin (Barc)* 2016;**147**:262–6. <https://doi.org/10.1016/j.medcli.2016.02.025>
24. Guyot P, Ades AE, Ouwens MJ, Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan–Meier survival curves. *BMC Med Res Methodol* 2012;**12**:9. <https://doi.org/10.1186/1471-2288-12-9>
25. Irvine AF, Waise S, Green EW, Stuart B. A non-linear optimisation method to extract summary statistics from Kaplan–Meier survival plots using the published P value. *BMC Med Res Methodol* 2020;**20**:269. <https://doi.org/10.1186/s12874-020-01092-x>
26. Bland JM, Altman DG. The logrank test. *BMJ* 2004;**328**:1073. <https://doi.org/10.1136/bmj.328.7447.1073>
27. Higgins JP, Savovic J, Page MJ, Sterne JA, ROB2 Development Group. *Revised Cochrane Risk-of-Bias tool for Randomized Trials (RoB 2)*. 2019. URL: <https://sites.google.com/site/riskofbiastool/welcome/rob-2-0-tool/current-version-of-rob-2> (accessed 27 November 2019).
28. Dias S, Ades AE, Welton NJ, Jansen JP, Sutton AJ. *Network Meta-analysis for Decision Making*. Hoboken, NJ: Wiley; 2018.
29. Dias S, Sutton AJ, Ades AE, Welton NJ. Evidence synthesis for decision making 2: a generalized linear modeling framework for pairwise and network meta-analysis of randomized controlled trials. *Med Decis Making* 2013;**33**:607–17.
30. Dias S, Welton NJ, Sutton AJ, Ades AE. *NICE DSU Technical Support Document 2: A Generalised Linear Modelling Framework for Pairwise and Network Meta-analysis of Randomised Controlled Trials*. London: NICE; 2014.
31. Warn DE, Thompson SG, Spiegelhalter DJ. Bayesian random effects meta-analysis of trials with binary outcomes: methods for the absolute risk difference and relative risk scales. *Stat Med* 2002;**21**:1601–23. <https://doi.org/10.1002/sim.1189>
32. van Valkenhoef G, Kuiper J. *gemtc: Network Meta-Analysis Using Bayesian Methods*. R package. CRAN; 2021.
33. Franchini A, Dias S, Ades A, Jansen J, Welton NJ. Accounting for correlation in mixed treatment comparisons with multi-arm trials. *Res Synth Methods* 2012;**3**:142–60.

34. Brooks SP, Gelman A. General methods for monitoring convergence of iterative simulations. *J Comput Graph Stat* 1998;**7**:434–55. <https://doi.org/10.1080/10618600.1998.10474787>
35. Röver C, Bender R, Dias S, Schmid CH, Schmidli H, Sturtz S, *et al.* On weakly informative prior distributions for the heterogeneity parameter in Bayesian random-effects meta-analysis. *Res Synth Methods* 2021;**12**:448–74. <https://doi.org/10.1002/jrsm.1475>
36. Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A. Bayesian measures of model complexity and fit. *J R Stat Soc Ser B Stat Methodol* 2002;**64**:583–616.
37. Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A. The deviance information criterion: 12 years on. *J R Stat Soc Ser B Stat Methodol* 2014;**76**:485–93.
38. Dias S, Welton NJ, Sutton AJ, Caldwell DM, Lu G, Ades AE. *NICE DSU Technical Support Document 4: Inconsistency in Networks of Evidence Based on Randomised Controlled Trials*. London: NICE; 2011.
39. Dias S, Welton NJ, Caldwell DM, Ades AE. Checking consistency in mixed treatment comparison meta-analysis. *Stat Med* 2010;**29**:932–44.
40. Phillippo DM, Dias S, Ades AE, Didelez V, Welton NJ. Sensitivity of treatment recommendations to bias in network meta-analysis. *J R Stat Soc Ser A Stat Soc* 2018;**181**:843–67. <https://doi.org/10.1111/rssa.12341>
41. Phillippo DM, Dias S, Welton NJ, Caldwell DM, Taske N, Ades AE. Threshold analysis as an alternative to GRADE for assessing confidence in guideline recommendations based on network meta-analyses. *Ann Intern Med* 2019;**170**:538–46. <https://doi.org/10.7326/M18-3542>
42. Hausner E, Metzendorf MI, Richter B, Lotz F, Waffenschmidt S. Study filters for non-randomized studies of interventions consistently lacked sensitivity upon external validation. *BMC Med Res Methodol* 2018;**18**:171. <https://doi.org/10.1186/s12874-018-0625-4>
43. Aikata H, Shirakawa H, Takaki S, Uka K, Miki D, Yamashina K. Radiofrequency ablation combined with transcatheter arterial chemoembolization for small hepatocellular carcinomas. *Hepatology* 2006;**44**:494A.
44. Brunello F, Veltri A, Carucci P, Pagano E, Ciccone G, Moretto P, *et al.* Radiofrequency ablation versus ethanol injection for early hepatocellular carcinoma: a randomized controlled trial. *Scand J Gastroenterol* 2008;**43**:727–35. <https://dx.doi.org/10.1080/00365520701885481>
45. Chen K, Chen G, Wang H, Li H, Xiao J, Duan X, *et al.* Increased survival in hepatocellular carcinoma with iodine-125 implantation plus radiofrequency ablation: a prospective randomized controlled trial. *J Hepatol* 2014;**61**:1304–11. <https://dx.doi.org/10.1016/j.jhep.2014.07.026>
46. Fang Y, Chen W, Liang X, Li D, Lou H, Chen R, *et al.* Comparison of long-term effectiveness and complications of radiofrequency ablation with hepatectomy for small hepatocellular carcinoma. *J Gastroenterol Hepatol* 2014;**29**:193–200. <https://dx.doi.org/10.1111/jgh.12441>
47. Gan YH, Yie SL, Ren ZG, Xia JL, Zhang BH, Wang YH, *et al.* Prospective randomized trial of RFA and chemotherapy for unresectable small hepatocellular carcinoma. *Zhonghua Zhong Liu Za Zhi* 2004;**26**:496–8.
48. Giorgio A, Di Sarno A, De Stefano G, Scognamiglio U, Farella N, Mariniello A, *et al.* Percutaneous radiofrequency ablation of hepatocellular carcinoma compared to percutaneous ethanol injection in treatment of cirrhotic patients: an Italian randomized controlled trial. *Anticancer Res* 2011;**31**:2291–5.
49. Huang GT, Lee PH, Tsang YM, Lai MY, Yang PM, Hu RH, *et al.* Percutaneous ethanol injection versus surgical resection for the treatment of small hepatocellular carcinoma: a prospective study. *Ann Surg* 2005;**242**:36–42.

50. Izumi N, Hasegawa K, Nishioka Y, Takayama T, Yamanaka N, Kudo M, *et al.* A multicenter randomized controlled trial to evaluate the efficacy of surgery vs. radiofrequency ablation for small hepatocellular carcinoma (SURF trial). *J Clin Oncol* 2019;**37**(15\_suppl):4002.
51. Kim TH, Koh YH, Kim BH, Kim MJ, Lee JH, Park B, *et al.* Proton beam radiotherapy vs. radiofrequency ablation for recurrent hepatocellular carcinoma: a randomized phase III trial. *J Hepatol* 2020;**74**:603–12. <https://dx.doi.org/10.1016/j.jhep.2020.09.026>
52. Koda M, Murawaki Y, Mitsuda A, Oyama K, Okamoto K, Idobe Y, *et al.* Combination therapy with transcatheter arterial chemoembolization and percutaneous ethanol injection compared with percutaneous ethanol injection alone for patients with small hepatocellular carcinoma: a randomized control study. *Cancer* 2001;**92**:1516–24.
53. Lin SM, Lin CJ, Lin CC, Hsu CW, Chen YC. Randomised controlled trial comparing percutaneous radiofrequency thermal ablation, percutaneous ethanol injection, and percutaneous acetic acid injection to treat hepatocellular carcinoma of 3 cm or less. *Gut* 2005;**54**:1151–6.
54. Shibata T, Isoda H, Hirokawa Y, Arizono S, Shimada K, Togashi K. Small hepatocellular carcinoma: is radiofrequency ablation combined with transcatheter arterial chemoembolization more effective than radiofrequency ablation alone for treatment? *Radiology* 2009;**252**:905–13. <https://dx.doi.org/10.1148/radiol.2523081676>
55. Shiina S, Teratani T, Obi S, Sato S, Tateishi R, Fujishima T, *et al.* A randomized controlled trial of radiofrequency ablation with ethanol injection for small hepatocellular carcinoma. *Gastroenterology* 2005;**129**:122–30.
56. Yan SY, Zhang Y, Sun C, Cao HX, Li GM, Wang YQ, *et al.* The clinical effect and relevant mechanism of combined sorafenib and radiofrequency ablation in the treatment of early small hepatocellular carcinoma. *Oncol Lett* 2016;**12**:951–5.
57. Zou D, Pan D, Deng L, Mao W, Zhang F, Miao J. Comparison of ultrasound-guided laser ablation and radiofrequency ablation in the treatment of small hepatocellular carcinoma. *Int J Clin Exp Med* 2017;**10**:9562–8.
58. Feng K, Yan J, Li X, Xia F, Ma K, Wang S, *et al.* A randomized controlled trial of radiofrequency ablation and surgical resection in the treatment of small hepatocellular carcinoma. *J Hepatol* 2012;**57**:794–802. <https://dx.doi.org/10.1016/j.jhep.2012.05.007>
59. Lin SM, Lin CJ, Lin CC, Hsu CW, Chen YC. Radiofrequency ablation improves prognosis compared with ethanol injection for hepatocellular carcinoma < or =4 cm. *Gastroenterology* 2004 **127**:1714–23.
60. Mizuki A, Tatemichi M, Tsukada N, Nagamatsu R, Kawaguchi M, Itoshima T, *et al.* Addition of transcatheter arterial chemoembolization decreased local recurrence but had no survival benefit to percutaneous ethanol injection therapy for patients with small hepatocellular carcinoma: a multicenter randomized control study. *Oncol Lett* 2010;**1**:855–9.
61. Orlacchio A, Bolacchi F, Chegai F, Bergamini A, Costanzo E, Del Giudice C, *et al.* Comparative evaluation of percutaneous laser and radiofrequency ablation in patients with HCC smaller than 4 cm. *Radiol Med* 2014;**119**:298–308. <https://dx.doi.org/10.1007/s11547-013-0339-y>
62. Shibata T, Iimuro Y, Yamamoto Y, Maetani Y, Ametani F, Itoh K, *et al.* Small hepatocellular carcinoma: comparison of radio-frequency ablation and percutaneous microwave coagulation therapy. *Radiology* 2002;**223**:331–7.
63. Vietti V, Duran R, Guiu B, Cercueil JP, Aube C, Digkila A, *et al.* Efficacy of microwave ablation versus radiofrequency ablation for the treatment of hepatocellular carcinoma in patients with chronic liver disease: a randomised controlled phase 2 trial. *Lancet Gastroenterol Hepatol* 2018;**3**:317–25. [https://dx.doi.org/10.1016/S2468-1253\(18\)30029-3](https://dx.doi.org/10.1016/S2468-1253(18)30029-3)

64. Abdelaziz A, Elbaz T, Shousha HI, Mahmoud S, Ibrahim M, Abdelmaksoud A, *et al.* Efficacy and survival analysis of percutaneous radiofrequency versus microwave ablation for hepatocellular carcinoma: an Egyptian multidisciplinary clinic experience. *Surg Endosc* 2014;**28**:3429–34. <https://dx.doi.org/10.1007/s00464-014-3617-4>
65. Azab M, Zaki S, El-Shetey AG, Abdel-Moty MF, Alnoomani NM, Gomaa AA, *et al.* Radiofrequency ablation combined with percutaneous ethanol injection in patients with hepatocellular carcinoma. *Arab J Gastroenterol* 2011;**12**:113–8. <https://dx.doi.org/10.1016/j.ajg.2011.07.005>
66. Chen MS, Zhang YJ, Li JQ, Liang HH, Zhang YQ, Zheng Y. Randomized clinical trial of percutaneous radiofrequency ablation plus absolute ethanol injection compared with radiofrequency ablation alone for small hepatocellular carcinoma. *Zhonghua Zhong Liu Za Zhi* 2005;**27**:623–5.
67. Chen MS, Li JQ, Liang HH, Lin XJ, Guo RP, Zheng Y, *et al.* Comparison of effects of percutaneous radiofrequency ablation and surgical resection on small hepatocellular carcinoma. *Zhonghua Yi Xue Za Zhi* 2005;**85**:80–3.
68. Chen MS, Li JQ, Zheng Y, Guo RP, Liang HH, Zhang YQ, *et al.* A prospective randomized trial comparing percutaneous local ablative therapy and partial hepatectomy for small hepatocellular carcinoma. *Ann Surg* 2006;**243**:321–8.
69. Huang J, Yan L, Cheng Z, Wu H, Du L, Wang J, *et al.* A randomized trial comparing radiofrequency ablation and surgical resection for HCC conforming to the Milan criteria. *Ann Surg* 2010;**252**:903–12. <https://dx.doi.org/10.1097/SLA.0b013e3181efc656>
70. Huo TI, Huang YH, Wu JC, Chiang JH, Lee PC, Chang FY, *et al.* Sequential transarterial chemoembolization and percutaneous acetic acid injection therapy versus repeated percutaneous acetic acid injection for unresectable hepatocellular carcinoma: a prospective study. *Ann Oncol* 2003;**14**:1648–53.
71. Lencioni RA, Allgaier HP, Cioni D, Olschewski M, Deibert P, Crocetti L, *et al.* Small hepatocellular carcinoma in cirrhosis: randomized comparison of radio-frequency thermal ablation versus percutaneous ethanol injection. *Radiology* 2003;**228**:235–40.
72. Liu H, Wang ZG, Fu SY, Li AJ, Pan ZY, Zhou WP, *et al.* Randomized clinical trial of chemoembolization plus radiofrequency ablation versus partial hepatectomy for hepatocellular carcinoma within the Milan criteria. *Br J Surg* 2016;**103**:348–56. <https://dx.doi.org/10.1002/bjs.10061>
73. Ng KK, Chok KS, Chan AC, Cheung TT, Wong TC, Lo CM. Long-term analysis of a prospective randomized trial of hepatic resection versus radiofrequency ablation for early stage hepatocellular carcinoma. *J Hepatobiliary Pancreat Sci* 2017;**24**:A88.
74. Peng ZW, Zhang YJ, Liang HH, Lin XJ, Guo RP, Chen MS. Recurrent hepatocellular carcinoma treated with sequential transcatheter arterial chemoembolization and RF ablation versus RF ablation alone: a prospective randomized trial. *Radiology* 2012;**262**:689–700. <https://doi.org/10.1148/radiol.11110637>
75. Xia Y, Li J, Liu G, Wang K, Qian G, Lu Z, *et al.* Long-term effects of repeat hepatectomy vs percutaneous radiofrequency ablation among patients with recurrent hepatocellular carcinoma: a randomized clinical trial. *JAMA Oncol* 2020;**6**:255–63. <https://dx.doi.org/10.1001/jamaoncol.2019.4477>
76. Zhang YJ, Liang HH, Chen MS, Guo RP, Li JQ, Zheng Y, *et al.* Hepatocellular carcinoma treated with radiofrequency ablation with or without ethanol injection: a prospective randomized trial. *Radiology* 2007;**244**:599–607.
77. Ferrari FS, Megliola A, Scorzelli A, Stella A, Vigni F, Drudi FM, *et al.* Treatment of small HCC through radiofrequency ablation and laser ablation. Comparison of techniques and long-term results. *Radiol Med* 2007;**112**:377–93.

78. Bian H, Zheng JS, Nan G, Li R, Chen C, Hu CX, *et al.* Randomized trial of [<sup>131</sup>I] metuximab in treatment of hepatocellular carcinoma after percutaneous radiofrequency ablation. *J Natl Cancer Inst* 2014;**106**:dju239. <https://dx.doi.org/10.1093/jnci/dju239>
79. Ng KK, Chok KS, Chan AC, Cheung TT, Wong TC, Fung JY, *et al.* Randomized clinical trial of hepatic resection versus radiofrequency ablation for early-stage hepatocellular carcinoma. *Br J Surg* 2017;**104**:1775–84. <https://dx.doi.org/10.1002/bjs.10677>
80. Zhu F, Chang Q, Duan S, Leng W. Efficacy and safety of radiofrequency ablation versus laparoscopic hepatectomy for small hepatocellular carcinoma: a protocol for a randomized controlled trial. *Medicine (Baltim)* 2021;**100**:e23678. <https://dx.doi.org/10.1097/MD.00000000000023678>
81. Schoenfeld D. Partial residuals for the proportional hazards regression model. *Biometrika* 1982;**69**:239–41. <https://doi.org/10.1093/biomet/69.1.239>
82. Huo TI, Huang YH, Wu JC, Lee PC, Chang FY, Lee SD. Comparison of percutaneous acetic acid injection and percutaneous ethanol injection for hepatocellular carcinoma in cirrhotic patients: a prospective study. *Scand J Gastroenterol* 2003;**38**:770–8.
83. Cheung TT, Fan ST, Chu FS, Jenkins CR, Chok KS, Tsang SH, *et al.* Survival analysis of high-intensity focused ultrasound ablation in patients with small hepatocellular carcinoma. *HPB (Oxford)* 2013;**15**:567–73. <https://dx.doi.org/10.1111/hpb.12025>
84. Choi MS, Hong SN, Lee JH, Koh KC, Paik SW, Yoo BC, *et al.* Radiofrequency ablation vs. surgical resection in the treatment of small hepatocellular carcinoma – a comparative study. *J Hepatol* 2004;**40**:75–6. [https://doi.org/10.1016/s0168-8278\(04\)90238-3](https://doi.org/10.1016/s0168-8278(04)90238-3)
85. Elgendi A, Elshafei M, Abdel-Azizi F, Bedewy E. Intraoperative ablation for small HCC not amenable for percutaneous radio frequency ablation in Child A cirrhotic patients. *HPB (Oxford)* 2014;**16**:227–8.
86. Horigome H, Nomura T, Nakao H, Saso K, Takahashi Y, Akita S, *et al.* Treatment of solitary small hepatocellular carcinoma: consideration of hepatic functional reserve and mode of recurrence. *Hepatogastroenterology* 2000;**47**:507–11.
87. Huang G, Chen X, Lau WY, Shen F, Wang RY, Yuan SX, *et al.* Quality of life after surgical resection compared with radiofrequency ablation for small hepatocellular carcinomas. *Br J Surg* 2014;**101**:1006–15. <https://dx.doi.org/10.1002/bjs.9539>
88. Peng ZW, Chen MS, Lin XJ, Zhang YJ. A prospective study on radiofrequency ablation versus surgical resection for small solitary hepatocellular carcinoma. *Hepatol Int* 2010;**4**:304. <http://dx.doi.org/10.1007/s12072-010-9169-3>
89. Qian GJ, Wang N, Shen Q, Sheng YH, Zhao JQ, Kuang M, *et al.* Efficacy of microwave versus radiofrequency ablation for treatment of small hepatocellular carcinoma: experimental and clinical studies. *Eur Radiol* 2012;**22**:1983–90. <https://dx.doi.org/10.1007/s00330-012-2442-1>
90. Tateishi R, Hasegawa K, Kawaguchi Y, Takayama T, Izumi N, Yamanaka N, *et al.* A multicenter nonrandomized controlled trial to evaluate the efficacy of surgery versus radiofrequency ablation for small hepatocellular carcinoma (SURF cohort trial). *J Clin Oncol* 2020;**38**:4581. [http://dx.doi.org/10.1200/JCO.2020.38.15\\_suppl.4581](http://dx.doi.org/10.1200/JCO.2020.38.15_suppl.4581)
91. Elgendi AM, Elshafey M, Bdeawey E. Intraoperative radiofrequency ablation versus surgical resection in solitary small HCC. *HPB (Oxford)* 2015;**17**:25. [http://dx.doi.org/10.1111/hpb.12399\\_8/abstract](http://dx.doi.org/10.1111/hpb.12399_8/abstract)
92. Du J. The curative effect of percutaneous RFA and radical resection to small hepatocellular carcinoma. *Clin J Med Offic* 2012;**40**:570–2.
93. Barabino M, Santambrogio R, Chiang J, Meloni F, Bruno S, Zuin M, *et al.* Laparoscopic thermal ablation of hepatocellular carcinoma unsuitable for percutaneous treatments or hepatic

- resection: a comparative study. *Dig Liver Dis* 2016;**48**:e47. <http://dx.doi.org/10.1016/j.dld.2015.12.109>
94. Sugimoto K, Kakimi K, Takeuchi H, Fujieda N, Saito K, Sato E, *et al*. Irreversible electroporation versus radiofrequency ablation: comparison of systemic immune responses in patients with hepatocellular carcinoma. *J Vasc Interv Radiol* 2019;**30**:845–53. <https://dx.doi.org/10.1016/j.jvir.2019.03.002>
  95. Ei S, Hibi T, Tanabe M, Itano O, Shinoda M, Kitago M, *et al*. Cryoablation provides superior local control of primary hepatocellular carcinomas of >2 cm compared with radiofrequency ablation and microwave coagulation therapy: an underestimated tool in the toolbox. *Ann Surg Oncol* 2015;**22**:1294–300. <https://dx.doi.org/10.1245/s10434-014-4114-7>
  96. Harada N, Shirabe K, Maeda T, Kayashima H, Takaki S, Maehara Y. Comparison of the outcomes of patients with hepatocellular carcinoma and portal hypertension after liver resection versus radiofrequency ablation. *World J Surg* 2016;**40**:1709–19. <https://dx.doi.org/10.1007/s00268-016-3465-6>
  97. National Institute for Health and Care Excellence. *Guide to the Methods of Technology Appraisal*. London: NICE; 2013.
  98. Canadian Agency for Drugs and Technologies in Health. *Strings Attached: CADTH's Database Search Filters - Narrow Economic Filter - OVID MEDLINE, Embase*. CADTH; 2021. URL: <https://www.cadth.ca/strings-attached-cadths-database-search-filters#narrow> (accessed 13 May 2021).
  99. Cucchetti A, Piscaglia F, Cescon M, Colecchia A, Bolondi L, Pinna AD. Cost-effectiveness of hepatic resection versus percutaneous ablation for hepatocellular carcinoma within the Milan criteria. *J Hepatol* 2013;**58**:S112–3. <http://dx.doi.org/10.1016/S0168-8278%2813%2960267-6>
  100. Cucchetti A, Piscaglia F, Cescon M, Colecchia A, Ercolani G, Bolondi L, *et al*. Cost-effectiveness of hepatic resection versus percutaneous radiofrequency ablation for early hepatocellular carcinoma. *J Hepatol* 2013;**59**:300–7. <http://dx.doi.org/10.1016/j.jhep.2013.04.009>
  101. Cucchetti A, Piscaglia F, Cescon M, Colecchia A, Ercolani G, Bolondi L, *et al*. Radiofrequency ablation versus hepatic resection for early hepatocellular carcinoma: a cost-effectiveness perspective. *Dig Liver Dis* 2013;**45**:S5.
  102. Sarasin FP, Majno PE, Llovet JM, Bruix J, Mentha G, Hadengue A. Living donor liver transplantation for early hepatocellular carcinoma: a life-expectancy and cost-effectiveness perspective. *Hepatology* 2001;**33**:1073–9.
  103. Lai Y, Li K, Li J, Liu SX. Cost-effectiveness of navigated radiofrequency ablation for hepatocellular carcinoma in China. *Int J Technol Assess Health Care* 2014;**30**:400–8. <https://dx.doi.org/10.1017/S0266462314000452>
  104. Spolverato G, Vitale A, Bagante F, Connolly R, Pawlik TM. Liver resection for breast cancer liver metastases: a cost-utility analysis. *Ann Surg* 2017;**265**:792–9. <https://dx.doi.org/10.1097/SLA.0000000000001715>
  105. Spolverato G, Vitale A, Ejaz A, Kim Y, Maithel SK, Cosgrove DP, *et al*. The relative net health benefit of liver resection, ablation, and transplantation for early hepatocellular carcinoma. *World J Surg* 2015;**39**:1474–84. <https://dx.doi.org/10.1007/s00268-015-2987-7>
  106. Naugler WE, Sonnenberg A. Survival and cost-effectiveness analysis of competing strategies in the management of small hepatocellular carcinoma. *Liver Transpl* 2010;**16**:1186–94. <https://dx.doi.org/10.1002/lt.22129>

107. Rostambeigi N, Dekarske A, Austin E, Golzarian J, Cressman E. Simulation study on cost-effectiveness of radioembolization compared with trans-arterial chemoembolization for hepatocellular carcinoma. *J Vasc Interv Radiol* 2014;**25**:S104. <http://dx.doi.org/10.1016/j.jvir.2013.12.292>
108. Rostambeigi N, Dekarske AS, Austin EE, Golzarian J, Cressman EN. Cost effectiveness of radioembolization compared with conventional transarterial chemoembolization for treatment of hepatocellular carcinoma. *J Vasc Interv Radiol* 2014;**25**:1075–84. <https://dx.doi.org/10.1016/j.jvir.2014.04.014>
109. Lim KC, Wang VW, Siddiqui FJ, Shi L, Chan ES, Oh HC, *et al*. Cost-effectiveness analysis of liver resection versus transplantation for early hepatocellular carcinoma within the Milan criteria. *Hepatology* 2015;**61**:227–37. <https://dx.doi.org/10.1002/hep.27135>
110. McLernon DJ, Dillon J, Donnan PT. Health-state utilities in liver disease: a systematic review. *Med Decis Making* 2008;**28**:582–92. <https://doi.org/10.1177/0272989x08315240>
111. Cho SK, Hay JW, Barzi A. Cost-effectiveness analysis of regorafenib and TAS-102 in refractory metastatic colorectal cancer in the United States. *Clin Colorectal Cancer* 2018;**17**:e751–e61. <https://dx.doi.org/10.1016/j.clcc.2018.08.003>
112. National Health Service. *National Schedule of NHS Costs*. 2019. URL: <https://www.england.nhs.uk/national-cost-collection/> (accessed 1 December 2021).
113. Curtis LA, Burns A. *Unit Costs of Health and Social Care 2020*. Canterbury: Personal Social Services Research Unit (PSSRU), University of Kent; 2020.
114. Joint Formulary Committee. *British National Formulary*. NICE; 2021. URL: <https://bnf.nice.org.uk/> (accessed 1 December 2021).
115. An C, Li WZ, Huang ZM, Yu XL, Han YZ, Liu FY, *et al*. Small single perivascular hepatocellular carcinoma: comparisons of radiofrequency ablation and microwave ablation by using propensity score analysis. *Eur Radiol* 2021;**31**:4764–73. <https://doi.org/10.1007/s00330-020-07571-5>
116. Chang W, Lee JM, Lee DH, Kim YJ, Yoon JH, Han JK. Comparison of switching bipolar ablation with multiple cooled wet electrodes and switching monopolar ablation with separable clustered electrode in treatment of small hepatocellular carcinoma: a randomized controlled trial. *PLOS ONE* 2018;**13**:1–16. <http://dx.doi.org/10.1371/journal.pone.0192173>
117. Cho YK. A comparison of surgical resection and radiofrequency ablation for the treatment of single small hepatocellular carcinoma  $\leq 2$  cm. *Hepatology* 2014;**59**:1653. <http://dx.doi.org/10.1002/hep.26692>
118. Chong C, Fong A, Cheung S, Wong J, Lee KF, Lai P. Microwave versus radiofrequency ablation for hepatocellular carcinoma: a randomized controlled trial. *Hepatol Int* 2017;**11**:S237. <https://doi.org/10.1007/s12072-016-9783-9>
119. Chong C, Fong A, Chu C, Yu S, Fung A, Lok HT, *et al*. Microwave versus radiofrequency ablation for hepatocellular carcinoma: a randomized controlled trial. *J Hepatobiliary Pancreat Sci* 2017;**24**:A194.
120. Chong CCN, Lee KF, Fong AKW, Chu CCM, Yu SCH, Fung AKY, *et al*. Microwave versus radiofrequency ablation for hepatocellular carcinoma: a prospective randomized controlled trial. *Liver Cancer* 2018;**7**:118–9. <https://doi.org/10.1159/000490877>
121. Chong CCN, Lee KF, Cheung SYS, Chu CCM, Fong AKW, Wong J, *et al*. Prospective double-blinded randomized controlled trial of microwave versus radiofrequency ablation for hepatocellular carcinoma (McRFA trial). *HPB (Oxford)* 2020;**22**:1121–7. <https://dx.doi.org/10.1016/j.hpb.2020.01.008>

122. Crocetti L. Radiofrequency ablation versus resection for small hepatocellular carcinoma: are randomized controlled trials still needed? *Radiology* 2018;**287**:473–5. <http://dx.doi.org/10.1148/radiol.2018172822>
123. Di Costanzo GG, Tortora R, D'Adamo G, Galeota Lanza A, Carannante N, Lampasi F, *et al.* Radiofrequency ablation versus laser ablation for the treatment of small hepatocellular carcinoma: a randomized controlled trial. *Dig Liver Dis* 2011;**43**:S74. [https://doi.org/10.1016/s1590-8658\(11\)60029-7](https://doi.org/10.1016/s1590-8658(11)60029-7)
124. Di Costanzo GG, Tortora R, D'Adamo G, Galeota Lanza A, Carannante N, Lampasi F, *et al.* Radiofrequency ablation versus laser ablation for the treatment of small hepatocellular carcinoma: a randomized controlled trial. *Dig Liver Dis* 2011;**43**:S140. [https://doi.org/10.1016/s1590-8658\(11\)60207-7](https://doi.org/10.1016/s1590-8658(11)60207-7)
125. Di Costanzo GG, Tortora R, D'Adamo G, Lanza AG, Carannante N, Lampasi F, *et al.* Radiofrequency ablation versus laser ablation for the treatment of small hepatocellular carcinoma: a randomized controlled trial. *J Hepatol* 2011;**54**:S254–S5.
126. Di Costanzo GG, Tortora R, D'Adamo G, Galeota Lanza A, Lampasi F, Addario L, *et al.* Radiofrequency ablation versus laser ablation for the treatment of small hepatocellular carcinoma: a randomized controlled trial. *Dig Liver Dis* 2013;**45**:S5.
127. Di Costanzo GG, Tortora R, D'Adamo G, De Luca M, Lampasi F, Addario L, *et al.* Radiofrequency ablation versus laser ablation for the treatment of small hepatocellular carcinoma in cirrhosis: a randomized trial. *J Gastroenterol Hepatol* 2015;**30**:559–65. <https://dx.doi.org/10.1111/jgh.12791>
128. Duan J, Yue H, Liu K, Wu M, Yang J. Percutaneous radiofrequency ablation versus repeat hepatectomy for recurrent hepatocellular carcinoma: retrospective randomized control study. *J Med Colleges PLA* 2011;**26**:316–23. <http://dx.doi.org/10.1016/S1000-1948%2812%2960027-6>
129. DuBay DA, Sandroussi C, Kachura JR, Ho CS, Beecroft JR, Vollmer CM, *et al.* Radiofrequency ablation of hepatocellular carcinoma as a bridge to liver transplantation. *HPB (Oxford)* 2011;**13**:24–32. <https://dx.doi.org/10.1111/j.1477-2574.2010.00228.x>
130. Fan J, He S, Zheng Y. Analyses of clinical efficacy of ultrasound-guided radiofrequency ablation in liver cancer adjacent to the gallbladder and its prognosis. *J BUON* 2019;**24**:2411–7.
131. Fang WQ, Li SP, Zhang CQ, Xu L, Shi M, Chen MS, *et al.* Prophylaxis and clinical treatment for surgical margin recurrence of small primary hepatocellular carcinoma. *Ai Zheng* 2005;**24**:834–6.
132. Ferrer Puchol M, la Parra C, Esteban E, Vano M, Forment M, Vera A, *et al.* Comparison of doxorubicin-eluting bead transarterial chemoembolization (DEB-TACE) with conventional transarterial chemoembolization (TACE) for the treatment of hepatocellular carcinoma. *Radiologia* 2011;**53**:246–53. <https://dx.doi.org/10.1016/j.rx.2010.07.010>
133. Filippiadis D, Mazioti A, Velonakis G, Tsochantzis A, Tosoratti N, Kelekis A, *et al.* Continuous versus pulsed microwave ablation in the liver: any difference in intraoperative pain scores? *Ann Gastroenterol* 2021;**34**:80–4. <https://dx.doi.org/10.20524/aog.2020.0557>
134. Fong AK, Chong CC, Cheung SY, Wong J, Lee KF, Lai PB. Preliminary results on microwave versus radiofrequency ablation for hepatocellular carcinoma: a randomized controlled trial. *Surg Pract* 2016;**20**:3. <https://doi.org/10.1111/1744-1633.12205>
135. Frangakis C, Geschwind JF, Kim D, Chen Y, Koteish A, Hong K, *et al.* Chemoembolization decreases drop-off risk of hepatocellular carcinoma patients on the liver transplant list. *Cardiovasc Intervent Radiol* 2010;**34**:1–8. <http://dx.doi.org/10.1007/s00270-010-0077-7>

136. Fukushima T, Ikeda K, Kawamura Y, Sorin Y, Hosaka T, Kobayashi M, *et al.* Randomized controlled trial comparing the efficacy of impedance control and temperature control of radiofrequency interstitial thermal ablation for treating small hepatocellular carcinoma. *Oncology (Huntingt)* 2015;**89**:47–52. <https://dx.doi.org/10.1159/000375166>
137. Gerunda GE, Neri D, Merenda R, Barbazza F, Zangrandi F, Meduri F, *et al.* Role of transarterial chemoembolization before liver resection for hepatocarcinoma. *Liver Transpl* 2000;**6**:619–26. <http://dx.doi.org/10.1053/jlts.2000.8312>
138. Giorgio A, De Stefano G, Di Sarno A, Scognamiglio U, Farella N, Sabino G, *et al.* Three cm or less single hepatocellular carcinoma: a randomized controlled trial comparing percutaneous ethanol injection versus percutaneous radiofrequency ablation. *Dig Liver Dis* 2010;**42**:S42. [https://doi.org/10.1016/s1590-8658\(10\)60479-3](https://doi.org/10.1016/s1590-8658(10)60479-3)
139. Giorgio A, De Stefano G, Di Sarno A, Scognamiglio U, Nunzia F, Sabino G, *et al.* Comparison between PEI and RF in the treatment of HCC in cirrhotic patients with 3 cm or less single nodule: a prospective randomized controlled trial. *J Hepatol* 2010;**52**:S91. [https://doi.org/10.1016/s0168-8278\(10\)60215-2](https://doi.org/10.1016/s0168-8278(10)60215-2)
140. Guo J, Sun XN, Huang M. Evaluation of the efficacy for alternated treatment on primary liver cancer by interventional therapy in combination with fractionated stereotactic conformal radiotherapy. *Chin J Clin Oncol* 2005;**32**:1418–20.
141. Guo WX, Sun JX, Cheng YQ, Shi J, Li N, Xue J, *et al.* Percutaneous radiofrequency ablation versus partial hepatectomy for small centrally located hepatocellular carcinoma. *World J Surg* 2013;**37**:602–7. <https://dx.doi.org/10.1007/s00268-012-1870-z>
142. Ha TY, Hwang S, Lee YJ, Kim KH, Ko GY, Gwon DH, *et al.* Absence of benefit of transcatheter arterial chemoembolization (TACE) in patients with resectable solitary hepatocellular carcinoma. *World J Surg* 2016;**40**:1200–10. <https://doi.org/10.1007/s00268-015-3373-1>
143. Hayes. *Radioactive Yttrium-90 Microspheres for Treatment of Primary Liver Cancer*. United States: Hayes; 2008.
144. He HJ, Lv MY, Wang LP. Value of contrast-enhanced ultrasonography in microwave ablation of liver cancer. *Shi Jie Hua Ren Xiao Hua Za Zhi* 2018;**26**:195–8. <http://dx.doi.org/10.11569/wcjd.v26.i3.195>
145. Hong SN, Lee SY, Choi MS, Lee JH, Koh KC, Paik SW, *et al.* Comparing the outcomes of radiofrequency ablation and surgery in patients with a single small hepatocellular carcinoma and well-preserved hepatic function. *J Clin Gastroenterol* 2005;**39**:247–52.
146. Hsiao CY, Hu RH, Ho CM, Wu YM, Lee PH, Ho MC. Surgical resection versus radiofrequency ablation for Barcelona Clinic Liver Cancer very early stage hepatocellular carcinoma: long-term results of a single-center study. *Am J Surg* 2020;**220**:958–64. <http://dx.doi.org/10.1016/j.amjsurg.2020.03.017>
147. Hsu KF, Chu CH, Chan DC, Yu JC, Shih ML, Hsieh HF, *et al.* Superselective transarterial chemoembolization vs hepatic resection for resectable early-stage hepatocellular carcinoma in patients with Child–Pugh class A liver function. *Eur J Radiol* 2012;**81**:466–71. <https://dx.doi.org/10.1016/j.ejrad.2010.12.058>
148. Huang XL, Yang HY, Liu BY, Guo YW, Zhan-Sen E. Ultrasound guided percutaneous ethanol injection and hyperthermal distilled-water injection in treatment for HCC. *Chin J Intervent Imag Therapy* 2009;**6**:309–13.
149. Huang ZM, Zuo MX, Gu YK, Gu HF, Lai CX, Zhang TQ, *et al.* Computed tomography-guided radiofrequency ablation combined with transarterial embolization assisted by a three-dimensional visualization ablation planning system for hepatocellular carcinoma in challenging locations: a preliminary study. *Abdom Radiol (NY)* 2020;**45**:1181–92. <https://dx.doi.org/10.1007/s00261-020-02426-5>

150. Hung HH, Chiou YY, Hsia CY, Su CW, Chou YH, Chiang JH, *et al.* Survival rates are comparable after radiofrequency ablation or surgery in patients with small hepatocellular carcinomas. *J Clin Gastroenterol Hepatol* 2011;**9**:79–86. <https://doi.org/10.1016/j.cgh.2010.08.018>
151. Hyun D, Cho SK, Shin SW, Park KB, Park HS, Choo SW, *et al.* Early stage hepatocellular carcinomas not feasible for ultrasound-guided radiofrequency ablation: comparison of transarterial chemoembolization alone and combined therapy with transarterial chemoembolization and radiofrequency ablation. *Cardiovasc Intervent Radiol* 2016;**39**:417–25. <https://dx.doi.org/10.1007/s00270-015-1194-0>
152. Iida H, Aihara T, Ikuta S, Yamanaka N. Comparative study of percutaneous radiofrequency ablation and hepatic resection for small, poorly differentiated hepatocellular carcinomas. *Hepatol Res* 2014;**44**:E156–62. <https://dx.doi.org/10.1111/hepr.12264>
153. Ikeda M, Okada S, Ueno H, Okusaka T, Kuriyama H. Radiofrequency ablation and percutaneous ethanol injection in patients with small hepatocellular carcinoma: a comparative study. *Jpn J Clin Oncol* 2001;**31**:322–6.
154. Imai K, Beppu T, Chikamoto A, Doi K, Okabe H, Hayashi H, *et al.* Comparison between hepatic resection and radiofrequency ablation as first-line treatment for solitary small-sized hepatocellular carcinoma of 3 cm or less. *Hepatol Res* 2013;**43**:853–64. <https://dx.doi.org/10.1111/hepr.12035>
155. Jiang L, Yan L, Wen T, Li B, Zeng Y, Yang J, *et al.* Comparison of outcomes of hepatic resection and radiofrequency ablation for hepatocellular carcinoma patients with multifocal tumors meeting the Barcelona-Clinic Liver Cancer stage A classification. *J Am Coll Surg* 2015;**221**:951–61. <http://dx.doi.org/10.1016/j.jamcollsurg.2015.08.009>
156. Jiang FQ, Lu W, Yang C, Du P, Ma JP, Yang J, *et al.* Curative effect of transcatheter arterial chemoembolization combined with radiofrequency ablation in treating hepatic cell carcinoma and its effect on serum markers. *Cancer Biomark* 2017;**20**:17–22. <http://dx.doi.org/10.3233/CBM-160508>
157. Jiang T, Zeng ZC, Yang P, Hu Y. Exploration of superior modality: safety and efficacy of hypofractionated image-guided intensity modulated radiation therapy in patients with unresectable but confined intrahepatic hepatocellular carcinoma. *Can J Gastroenterol Hepatol* 2017;**2017**:6267981. <https://dx.doi.org/10.1155/2017/6267981>
158. Kaibori M, Tanigawa N, Kariya S, Ikeda H, Nakahashi Y, Hirohara J, *et al.* A prospective randomized controlled trial of preoperative whole-liver chemolipiodolization for hepatocellular carcinoma. *Dig Dis Sci* 2012;**57**:1404–12. <https://dx.doi.org/10.1007/s10620-012-2029-3>
159. Kaibori M, Ishizaki M, Matsui K, Kwon AH. A prospective randomized controlled trial of preoperative whole-liver chemolipiodolization for hepatocellular carcinoma. *Hepatol Int* 2012;**6**:241. <https://doi.org/10.1007/s12072-011-9333-4>
160. Kayali Z, Slater J, Bush D. Randomized controlled trial for proton beam radiotherapy versus transarterial chemoembolization for the treatment of hepatocellular carcinoma; preliminary results. *J Hepatol* 2013;**58**:S111. [https://doi.org/10.1016/s0168-8278\(13\)60262-7](https://doi.org/10.1016/s0168-8278(13)60262-7)
161. Kim W, Cho SK, Shin SW, Hyun D, Lee MW, Rhim H. Combination therapy of transarterial chemoembolization (TACE) and radiofrequency ablation (RFA) for small hepatocellular carcinoma: comparison with TACE or RFA monotherapy. *Abdom Radiol (NY)* 2019;**44**:2283–92. <https://dx.doi.org/10.1007/s00261-019-01952-1>
162. Kim TH, Koh YH, Kim BH, Kim MJ, Lee JH, Park B, *et al.* Proton beam radiotherapy vs. radiofrequency ablation for recurrent hepatocellular carcinoma: a randomized phase III trial. *J Hepatol* 2021;**74**:603–12. <https://doi.org/10.1016/j.jhep.2020.09.026>

163. Kitamoto M, Imagawa M, Yamada H, Watanabe C, Sumioka M, Satoh O, *et al.* Radiofrequency ablation in the treatment of small hepatocellular carcinomas: comparison of the radiofrequency effect with and without chemoembolization. *AJR Am J Roentgenol* 2003;**181**:997–1003. <http://dx.doi.org/10.2214/ajr.181.4.1810997>
164. Kiyoshi H, Norihiro K, Shuichiro S, Ryosuke T, Masatoshi M. Surgery versus radiofrequency ablation for small hepatocellular carcinoma: start of a randomized controlled trial (SURF trial). *Hepatol Res* 2010;**40**:851–2. <http://dx.doi.org/10.1111/j.1872-034X.2010.00696.x>
165. Kobayashi M, Ikeda K, Kawamura Y, Hosaka T, Sezaki H, Yatsuji H, *et al.* Randomized controlled trial for the efficacy of hepatic arterial occlusion during radiofrequency ablation for small hepatocellular carcinoma – direct ablative effects and a long-term outcome. *Liver Int* 2007;**27**:353–9.
166. Koda M, Murawaki Y, Kawasaki H. The combination therapy of transcatheter arterial embolization and percutaneous ethanol injection for small hepatocellular carcinoma: randomized controlled study. *J Hepatol* 2000;**32**:164.
167. Koh PS, Chan AC, Cheung TT, Chok KS, Dai WC, Poon RT, *et al.* Efficacy of radiofrequency ablation compared with transarterial chemoembolization for the treatment of recurrent hepatocellular carcinoma: a comparative survival analysis. *HPB (Oxford)* 2015;**16**:16. <https://dx.doi.org/10.1111/hpb.12495>
168. Kong QF, Jiao JB, Chen QQ, Li L, Wang DG, Lv B. Comparative effectiveness of radiofrequency ablation with or without transarterial chemoembolization for hepatocellular carcinoma. *Tumour Biol* 2014;**35**:2655–9.
169. Lai C, Jin RA, Liang X, Cai XJ. Comparison of laparoscopic hepatectomy, percutaneous radiofrequency ablation and open hepatectomy in the treatment of small hepatocellular carcinoma. *J Zhejiang Univ Sci B* 2016;**17**:236–46. <https://dx.doi.org/10.1631/jzus.B1500322>
170. Lambert B, Dhondt E, Hermie L, Verhelst X, Defreyne L. Transarterial radioembolization versus drug-eluting beads chemoembolization for treatment of inoperable early and intermediate hepatocellular carcinoma: interim results of the randomized controlled TRACE trial. *Eur J Nucl Med Mol Imaging* 2020;**47**:S57–S8.
171. Lee J. Hypofractionated radiotherapy for small-sized hepatocellular carcinoma as salvage therapy: sustained local control and safety. *Int J Radiat Oncol Biol Phys* 2009;**75**:S263. <https://doi.org/10.1016/j.ijrobp.2009.07.606>
172. Lee HW, Lee JM, Yoon JH, Kim YJ, Park JW, Park SJ, *et al.* A prospective randomized study comparing radiofrequency ablation and hepatic resection for hepatocellular carcinoma. *Ann Surg Treat Res* 2018;**94**:74–82. <https://dx.doi.org/10.4174/astr.2018.94.2.74>
173. Lee H, Yoon CJ, Seong NJ, Jeong SH, Kim JW. Comparison of combined therapy using conventional chemoembolization and radiofrequency ablation versus conventional chemoembolization for ultrasound-invisible early-stage hepatocellular carcinoma (Barcelona Clinic Liver Cancer stage 0 or A). *Korean J Radiol* 2018;**19**:1130–9. <https://dx.doi.org/10.3348/kjr.2018.19.6.1130>
174. Li XY, Yang LZ. Analysis of transcatheter arterial chemoembolization combined with percutaneous ethanol injection for hepatoma. *Chin J Interv Imaging Ther* 2007;**4**:269–72.
175. Li J. Intraarterial chemotherapy combined with microwave ablation in treatment of hepatocellular carcinoma. *J Pract Oncol* 2011;**26**:386–9.
176. Li H, Hu Y, Li N, Zhou Y. Liver fibrosis and five year survival of hepatocellular cancer cases undergoing transcatheter arterial chemo embolization using small doses. *Asian Pac J Cancer Prev* 2012;**13**:1589–93.

177. Liao M, Zhong X, Zhang J, Liu Y, Zhu Z, Wu H, *et al.* Radiofrequency ablation using a 10-mm target margin for small hepatocellular carcinoma in patients with liver cirrhosis: a prospective randomized trial. *J Surg Oncol* 2017;**115**:971–9. <https://dx.doi.org/10.1002/jso.24607>
178. Lin SM, Lin CJ, Lin CC, Hsu CW, Chen YC. Erratum: randomised controlled trial comparing percutaneous radiofrequency thermal ablation, percutaneous ethanol injection, and percutaneous acetic acid injection to treat hepatocellular carcinoma of 3 cm or less (Gut (August 2005) 54 (1151–1156)). *Gut* 2005;**54**:1508. <http://dx.doi.org/10.1136/gut.2004.045203corr1>
179. Lin SM, Chu CM. Percutaneous tumor ablation or surgical resection for small hepatocellular carcinoma? *J Gastroenterol Hepatol* 2007;**22**:1561–4.
180. Lin SM, Lin CC, Chen WT, Chen YC, Hsu CW. Radiofrequency ablation for hepatocellular carcinoma: a prospective comparison of four radiofrequency devices. *J Vasc Interv Radiol* 2007;**18**:1118–25.
181. Liu YS, Chuang MT, Tsai YS, Tsai HM, Lin XZ. Nitroglycerine use in transcatheter arterial (chemo)embolization in patients with hepatocellular carcinoma and dual-energy CT assessment of Lipiodol retention. *Eur Radiol* 2012;**22**:2193–200. <https://dx.doi.org/10.1007/s00330-012-2484-4>
182. Liu B, Huang JW, Li Y, Hu BS, He X, Zhao W, *et al.* Single-agent versus combination doxorubicin-based transarterial chemoembolization in the treatment of hepatocellular carcinoma: a single-blind, randomized, phase II trial. *Oncology (Huntingt)* 2015;**89**:23–30. <https://dx.doi.org/10.1159/000371522>
183. Liu R, Sun M, Shi L, Yu H. The efficacy assessment of radiofrequency ablation treatment on primary hepatocellular carcinoma by contrast enhanced ultrasonography. *Int J Clin Exp Med* 2016;**9**:16218–23.
184. Liu Z, Wang X, Xing L, Pan Y, Huang S. The application comparison of contrast-enhanced ultrasound and contrast-enhanced computed tomography in radiofrequency ablation treatment for hepatocellular carcinoma. *Cancer Biother Radiopharm* 2019;**34**:621–5. <https://dx.doi.org/10.1089/cbr.2019.2845>
185. Liu E, Li B, Huang Y, Zhang H. *Clinical Study of the Beneficial and Harmful Effects of Transcatheter Arterial Chemoembolization Combined with Radiofrequency Ablation Versus Transcatheter Arterial Chemoembolization Alone for Primary Liver Cancer Inadults*. 2020. URL: [http://www.crd.york.ac.uk/PROSPERO/display\\_record.asp?ID=CRD42020207801](http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42020207801) (accessed 3 February 2021).
186. Lu MD, Kuang M, Liang LJ, Xie XY, Peng BG, Liu GJ, *et al.* Surgical resection versus percutaneous thermal ablation for early-stage hepatocellular carcinoma: a randomized clinical trial. *Zhonghua Yi Xue Za Zhi* 2006;**86**:801–5.
187. Lu LG, Hu BS, Li Y, Luo PF, Chen XM. Research of high frequency induced thermotherapy efficiency for treatment of hepatocellular carcinoma. *Zhonghua Yi Xue Za Zhi* 2008;**88**:1618–20.
188. Luo BM, Wen YL, Yang HY, Zhi H, Xiao XY, Ou B, *et al.* Percutaneous ethanol injection, radiofrequency and their combination in treatment of hepatocellular carcinoma. *World J Gastroenterol* 2005;**11**:6277–80.
189. Ma Y, Zhao C, Zhao H, Li H, Chen C, Xiang H, *et al.* Comparison of treatment efficacy and safety between drug-eluting bead transarterial chemoembolization with CalliSpheres microspheres and conventional transarterial chemoembolization as first-line treatment in hepatocellular carcinoma patients. *Am J Transl Res* 2019;**11**:7456–70.
190. Maeda S, Shibata J, Fujiyama S, Tanaka M, Noumaru S, Sato K, *et al.* Long-term follow-up of hepatic arterial chemoembolization with cisplatin suspended in iodized oil for hepatocellular carcinoma. *Hepatogastroenterology* 2003;**50**:809–13.

191. Masuda T, Beppu T, Ishiko T, Horino K, Toyama E, Tanaka H, *et al.* Local ablation therapy combined with liver resection for multiple hepatocellular carcinomas. *Gan To Kagaku Ryoho* 2007;**34**:2077–9.
192. Mbalisike EC, Vogl TJ, Zangos S, Eichler K, Balakrishnan P, Paul J. Image-guided microwave thermoablation of hepatic tumours using novel robotic guidance: an early experience. *Eur Radiol* 2015;**25**:454–62. <https://dx.doi.org/10.1007/s00330-014-3398-0>
193. Meniconi RL, Komatsu S, Perdigao F, Boelle PY, Soubrane O, Scatton O. Recurrent hepatocellular carcinoma: a Western strategy that emphasizes the impact of pathologic profile of the first resection. *Surgery* 2015;**157**:454–62. <https://dx.doi.org/10.1016/j.surg.2014.10.011>
194. Meyer T, Kirkwood A, Roughton M, Beare S, Tsochatzis E, Yu D, *et al.* A randomised phase II/III trial of 3-weekly cisplatin-based sequential transarterial chemoembolisation vs embolisation alone for hepatocellular carcinoma. *Br J Cancer* 2013;**108**:1252–9. <https://dx.doi.org/10.1038/bjc.2013.85>
195. Mohamed B, Emara M, Tai C. Percutaneous injection of ethanol and mitoxantrone versus radiofrequency ablation in the treatment of hepatocellular carcinoma. *Hepatology* 2018;**68**:862A. <https://doi.org/10.1002/hep.30257>
196. Mornex F, Wautot V, Kubas A, Serres A, Maillard E, Trepo C, *et al.* High-dose 3D-conformal radiation therapy (CRT): a new curative treatment for patients with small hepatocellular carcinomas (HCC). Mature results of a French phase II trial (RTF1). *EJC Suppl* 2007;**5**:271.
197. Mornex F, Wautot V, Kubas A, Maillard E, Trepo C, Merle P. High-dose 3D-conformal radiation therapy (CRT): a new curative treatment for patients with small hepatocellular carcinomas (HCC). Mature results of a French phase II trial (RTF1). *Ann Oncol* 2007;**18**:VII102–VII.
198. Murakami T, Ishimaru H, Sakamoto I, Uetani M, Matsuoka Y, Daikoku M, *et al.* Percutaneous radiofrequency ablation and transcatheter arterial chemoembolization for hypervascular hepatocellular carcinoma: rate and risk factors for local recurrence. *Cardiovasc Intervent Radiol* 2007;**30**:696–704.
199. Ni LD, Fu QC, Chen CW, Wang XJ, Zhou F. The comparative study on treatment for hepatocellular carcinoma with decompensated hepatocirrhosis by different interventional modes. *J Interv Radiol* 2007;**16**:249–52.
200. National Institute for Health Research. *Doxorubicin-eluting Beads (DC Bead, DC Bead M1 and Radiopaque DC Bead) for Hepatocellular Carcinoma*. England: National Institute for Health Research; 2016.
201. Nouse K, Oonishi A, Wakuta A, Kariyama K. Modified radiofrequency ablation for the treatment of hepatocellular carcinoma. *Hepatol Res* 2016;**46**:1158–61. <https://dx.doi.org/10.1111/hepr.12683>
202. Ohmoto K, Yamamoto S. Comparison between radiofrequency ablation and percutaneous microwave coagulation therapy for small hepatocellular carcinomas. *Clin Radiol* 2006;**61**:800–1. <http://dx.doi.org/10.1016/j.crad.2006.04.015>
203. Ohmoto K, Yoshioka N, Tomiyama Y, Shibata N, Kawase T, Yoshida K, *et al.* Comparison of therapeutic effects between radiofrequency ablation and percutaneous microwave coagulation therapy for small hepatocellular carcinomas. *JGH Open* 2009;**24**:223–7. <http://dx.doi.org/10.1111/j.1440-1746.2008.05596.x>
204. Olschewski M, Lencioni R, Allgeier H, Cioni D, Deibert P, Frings H, *et al.* A randomized comparison of radiofrequency thermal ablation and percutaneous ethanol injection for the treatment of small hepatocellular carcinoma (abstract). *Proc Am Soc Clin Oncol* 2002;**20**:126A.
205. Paik EK, Kim MS, Jang WI, Seo YS, Cho CK, Yoo HJ, *et al.* Benefits of stereotactic ablative radiotherapy combined with incomplete transcatheter arterial chemoembolization

- in hepatocellular carcinoma. *Radiat Oncol* 2016;**11**:22. <https://dx.doi.org/10.1186/s13014-016-0597-7>
206. Panaro F, Ramos J, Gallix B, Mercier G, Herrero A, Niampa H, *et al.* Hepatic artery complications following liver transplantation. Does preoperative chemoembolization impact the postoperative course? *Clin Transplant* 2014;**28**:598–605. <https://dx.doi.org/10.1111/ctr.12358>
207. Park SJ, Cho EJ, Lee JH, Yu SJ, Kim YJ, Yoon JH, *et al.* Switching monopolar no-touch radiofrequency ablation using octopus electrodes for small hepatocellular carcinoma: a randomized clinical trial. *Liver Cancer* 2020;**10**:72–81. <http://dx.doi.org/10.1159/000512338>
208. Peng ZW, Zhang YJ, Chen MS, Liang HH, Li JQ, Zhang YQ, *et al.* Risk factors of survival after percutaneous radiofrequency ablation of hepatocellular carcinoma. *Surg Oncol* 2008;**17**:23–31. <http://dx.doi.org/10.1016/j.suronc.2007.08.002>
209. Petrowsky H, Busuttil RW. Resection or ablation of small hepatocellular carcinoma: what is the better treatment? *J Hepatol* 2008;**49**:502–4. <http://dx.doi.org/10.1016/j.jhep.2008.07.018>
210. Pompili M, Saviano A, de Matthaeis N, Cucchetti A, Ardito F, Federico B, *et al.* Long-term effectiveness of resection and radiofrequency ablation for single hepatocellular carcinoma  $\leq 3$  cm. Results of a multicenter Italian survey. *J Hepatol* 2013;**59**:89–97. <https://doi.org/10.1016/j.jhep.2013.03.009>
211. Riaz A, Ryu RK, Kulik LM, Mulcahy MF, Lewandowski RJ, Minocha J, *et al.* Alpha-fetoprotein response after locoregional therapy for hepatocellular carcinoma: oncologic marker of radiologic response, progression, and survival. *J Clin Oncol* 2009;**27**:5734–42. <https://doi.org/10.1200/jco.2009.23.1282>
212. Roche AJ. Results and complications of selective chemoembolization in limited HCC. *Cardiovasc Intervent Radiol* 2002;**25**:S123–5.
213. Ryu T, Takami Y, Wada Y, Tateishi M, Matsushima H, Mikagi K, *et al.* Double- and triple-positive tumor markers predict early recurrence and poor survival in patients with hepatocellular carcinoma within the Milan criteria and Child–Pugh class A. *J Gastrointest Surg* 2017;**21**:957–66. <https://dx.doi.org/10.1007/s11605-017-3394-1>
214. Santambrogio R, Opocher E, Zuin M, Selmi C, Bertolini E, Costa M, *et al.* Surgical resection versus laparoscopic radiofrequency ablation in patients with hepatocellular carcinoma and Child–Pugh class A liver cirrhosis. *Ann Surg Oncol* 2009;**16**:3289–98. <https://dx.doi.org/10.1245/s10434-009-0678-z>
215. Shen H, Zhou S, Lou Y, Gao Y, Cao S, Wu D, *et al.* Microwave-assisted ablation improves the prognosis of patients with hepatocellular carcinoma undergoing liver resection. *Technol Cancer Res Treat* 2018;**17**:1533033818785980. <https://dx.doi.org/10.1177/1533033818785980>
216. Sherman M. What is the long-term efficacy of percutaneous ethanol injection for small hepatocellular carcinoma? *Nat Clin Pract Gastroenterol Hepatol* 2006;**3**:78–9.
217. Shi J, Sun Q, Wang Y, Jing X, Ding J, Yuan Q, *et al.* Comparison of microwave ablation and surgical resection for treatment of hepatocellular carcinomas conforming to Milan criteria. *J Gastroenterol Hepatol* 2014;**29**:1500–7. <https://dx.doi.org/10.1111/jgh.12572>
218. Shibata T, Shibata T, Maetani Y, Isoda H, Hiraoka M. Radiofrequency ablation for small hepatocellular carcinoma: prospective comparison of internally cooled electrode and expandable electrode. *Radiology* 2006;**238**:346–53.
219. Sun Q, Shi J, Ren C, Du Z, Shu G, Wang Y. Survival analysis following microwave ablation or surgical resection in patients with hepatocellular carcinoma conforming to the Milan criteria. *Oncol Lett* 2020;**19**:4066–76. <https://dx.doi.org/10.3892/ol.2020.11529>

220. Tashiro H, Aikata H, Waki K, Amano H, Oshita A, Kobayashi T, *et al.* Treatment strategy for early hepatocellular carcinomas: comparison of radiofrequency ablation with or without transcatheter arterial chemoembolization and surgical resection. *J Surg Oncol* 2011;**104**:3–9. <https://dx.doi.org/10.1002/jso.21745>
221. Toyoda H, Kumada T, Kaneoka Y, Osaki Y, Kimura T, Arimoto A, *et al.* Prognostic value of pretreatment levels of tumor markers for hepatocellular carcinoma on survival after curative treatment of patients with HCC. *J Hepatol* 2008;**49**:223–32. <https://dx.doi.org/10.1016/j.jhep.2008.04.013>
222. Tsai WL, Cheng JS, Lai KH, Lin CP, Lo GH, Hsu PI, *et al.* Clinical trial: percutaneous acetic acid injection vs. percutaneous ethanol injection for small hepatocellular carcinoma – a long-term follow-up study. *Aliment Pharmacol Ther* 2008;**28**:304–11.
223. Vivarelli M, Guglielmi A, Ruzzenente A, Cucchetti A, Bellusci R, Cordiano C, *et al.* Surgical resection versus percutaneous radiofrequency ablation in the treatment of hepatocellular carcinoma on cirrhotic liver. *Ann Surg* 2004;**240**:102–7.
224. Wang C, Wang H, Yang W, Hu K, Xie H, Bai W, *et al.* A multicenter randomized controlled trial of percutaneous cryoablation versus radiofrequency ablation therapy in patients with hepatocellular carcinoma. *Hepatology* 2014;**60**:282A–3A. <https://doi.org/10.1002/hep.27475>
225. Wang C, Wang H, Yang W, Hu K, Xie H, Hu KQ, *et al.* Multicenter randomized controlled trial of percutaneous cryoablation versus radiofrequency ablation in hepatocellular carcinoma. *Hepatology* 2015;**61**:1579–90. <https://dx.doi.org/10.1002/hep.27548>
226. Wu J, Song ZJ. Long-term efficacy of hepatic artery chemoembolization combined with cryosurgical for the treatment of primary liver cancer. *Int J Clin Exp Med* 2016;**9**:16797–801.
227. Xu F, Huang YQ, Li YS, Wu L, Yang JM. Is postoperative adjuvant transcatheter arterial chemoembolization necessary for small hepatocellular carcinoma patients: a randomized controlled trial. *Ti Erh Chun I Ta Hsueh Hsueh Pao* 2012;**33**:274–9. <http://dx.doi.org/10.3724/SP.J.1008.2012.00274>
228. Xu F, Huang YQ, Li YS, Wu L, Yang JM. Role of adjuvant TACE in prevention of early recurrence for HCC in BCLC stage A1 after radical hepatectomy, a RCT. *HPB (Oxford)* 2013;**15**:131–2. <https://doi.org/10.1111/hpb.12093>
229. Xu C, Lv PH, Huang XE, Wang SX, Sun L, Wang FA. Efficacy of transarterial chemoembolization combined with radiofrequency ablation in treatment of hepatocellular carcinoma. *Asian Pac J Cancer Prev* 2015;**16**:6159–62.
230. Xu J, Zhao Y. Comparison of percutaneous microwave ablation and laparoscopic resection in the prognosis of liver cancer. *Int J Clin Exp Pathol* 2015;**8**:11665–9.
231. Yamamoto J, Okada S, Shimada K, Okusaka T, Yamasaki S, Ueno H, *et al.* Treatment strategy for small hepatocellular carcinoma: comparison of long-term results after percutaneous ethanol injection therapy and surgical resection. *Hepatology* 2001;**34**:707–13.
232. Yamasaki T, Hamabe S, Saeki I, Harima Y, Yamaguchi Y, Uchida K, *et al.* A novel transcatheter arterial infusion chemotherapy using iodized oil and degradable starch microspheres for hepatocellular carcinoma: a prospective randomized trial. *J Gastroenterol* 2011;**46**:359–66. <http://dx.doi.org/10.1007/s00535-010-0306-5>
233. Yin L, Li H, Li AJ, Lau WY, Pan ZY, Lai EC, *et al.* Partial hepatectomy vs. transcatheter arterial chemoembolization for resectable multiple hepatocellular carcinoma beyond Milan criteria: a RCT. *J Hepatol* 2014;**61**:82–8. <https://dx.doi.org/10.1016/j.jhep.2014.03.012>
234. Yin L, Zhou W, Lau W, Lai E, Wu M. Partial hepatectomy versus transcatheter arterial chemoembolization for resectable multiple hepatocellular carcinoma beyond Milan criteria: a RCT. *HPB (Oxford)* 2015;**17**:11.

235. Yin X, Hua T, Liang C, Chen Z. Efficacy of re-resection versus radiofrequency ablation for recurrent Barcelona Clinic Liver Cancer stage 0/A hepatocellular carcinoma (HCC) after resection for primary HCC. *Transl Cancer Res* 2019;**8**:1035–45. <http://dx.doi.org/10.21037/tcr.2019.06.11>
236. Yu J, Liang P, Yu X, Cheng Z, Han Z, Liu F. Comparison of cooled-probe microwave and radiofrequency ablation treatment in incipient hepatocellular carcinoma: a phase III randomized controlled trial with 6-year follow-up. *J Clin Oncol* 2016;**34**:4068.
237. Yuan G, Zeng CL, Zhu DD, Shi XJ. Influences of RFA combined with TACE on the HIF-1alpha and EGR level of patients with primary hepatic carcinoma. *Eur Rev Med Pharmacol Sci* 2017;**21**:1738–45.
238. Yuen MF, Chan AO, Wong BC, Hui CK, Ooi GC, Tso WK, *et al*. Transarterial chemoembolization for inoperable, early stage hepatocellular carcinoma in patients with Child–Pugh grade A and B: results of a comparative study in 96 Chinese patients. *Am J Gastroenterol* 2003;**98**:1181–5.
239. Yun WK, Choi MS, Choi D, Rhim HC, Joh JW, Kim KH, *et al*. Superior long-term outcomes after surgery in Child–Pugh class A patients with single small hepatocellular carcinoma compared to radiofrequency ablation. *Hepatol Int* 2011;**5**:722–9. <http://dx.doi.org/10.1007/s12072-010-9237-8>
240. Zeng JY, Piao XH, Zou ZY, Yang QF, Qin ZL, Chen JB, *et al*. Cryoablation with drug-loaded bead embolization in the treatment of unresectable hepatocellular carcinoma: safety and efficacy analysis. *Oncotarget* 2018;**9**:7557–66. <http://dx.doi.org/10.18632/oncotarget.24029>
241. Zhang L, Wang N, Shen Q, Cheng W, Qian GJ. Therapeutic efficacy of percutaneous radiofrequency ablation versus microwave ablation for hepatocellular carcinoma. *PLOS ONE* 2013;**8**:e76119. <http://dx.doi.org/10.1371/journal.pone.0076119>
242. Zhang K, Jiang L, Jia Z, Zhang Y, He R, Ding Z, *et al*. Radiofrequency ablation plus devascularization is the preferred treatment of hepatocellular carcinoma with esophageal varices. *Dig Dis Sci* 2015;**60**:1490–501. <https://dx.doi.org/10.1007/s10620-014-3455-1>
243. Zhang B, Zhou CJ, Liu QH, Yang W, Liu YL. Percutaneous cryoablation versus hepatectomy for hepatocellular carcinoma treatment. *Int J Clin Exp Med* 2016;**9**:10354–63.
244. Zhao M, Wang JP, Li W, Huang ZL, Zhang FJ, Fan WJ, *et al*. Comparison of safety and efficacy for transcatheter arterial chemoembolization alone and plus radiofrequency ablation in the treatment of single branch portal vein tumor thrombus of hepatocellular carcinoma and their prognosis factors. *Zhonghua Yi Xue Za Zhi* 2011;**91**:1167–72.
245. Zhi, Luo BM, Wen YL, Yang HY, Ou B, Ma JH. Comparative study on treatment of small hepatocellular carcinoma by radiofrequency ablation and percutaneous ethanol injection. *Chin J Interv Imaging Ther* 2006;**3**:291–3.
246. Zhou P, Liu X, Li R, Nie W. Percutaneous coagulation therapy of hepatocellular carcinoma by combining microwave coagulation therapy and ethanol injection. *Eur J Radiol* 2009;**71**:338–42. <http://dx.doi.org/10.1016/j.ejrad.2008.04.010>
247. Zhou Z, Lei J, Li B, Yan L, Wang W, Wei Y, *et al*. Liver resection and radiofrequency ablation of very early hepatocellular carcinoma cases (single nodule <2 cm): a single-center study. *Eur J Gastroenterol Hepatol* 2014;**26**:339–44. <https://dx.doi.org/10.1097/MEG.000000000000012>
248. Zhou GH, Sun JH, Zhang YL, Zhou TY, Nie CH, Zhu TY, *et al*. Transcatheter embolization of hepatocellular carcinoma with epirubicin-loaded DC beads in Chinese patients. *Transl Cancer Res* 2019;**8**:279–89. <http://dx.doi.org/10.21037/tcr.2019.01.36>
249. Zhu W, Wang S, Zhang J. Combination treatment of unresectable hepatic carcinoma with percutaneous cryotherapy and ethanol injection therapy. *Technol Cancer Res Treat* 2007;**6**:564–7.

250. Zhu D, Yuan D, Wang Z, Chen S. Efficacy of drug-eluting bead transarterial chemoembolization (DEB-TACE) combined with radiofrequency ablation versus DEB-TACE alone in Chinese hepatocellular carcinoma patients. *Medicine (Baltim)* 2019;**98**:e15682. <https://dx.doi.org/10.1097/MD.00000000000015682>
251. Aikata H, Hiramatsu A, Tsuge M, Kawaoka T, Imamura M, Murakami E, et al. Stereotactic body radiation therapy combined with transcatheter arterial chemoembolization for small hepatocellular carcinoma. *Hepatology* 2018;**68**:544A–5A.
252. Al-Judaibi B, Dokus MK. Surgical resection vs radiofrequency ablation in older adults with early stage hepatocellular carcinoma: where do we stand? *Saudi J Gastroenterol* 2018;**24**:309–10. [https://dx.doi.org/10.4103/sjg.SJG\\_501\\_18](https://dx.doi.org/10.4103/sjg.SJG_501_18)
253. Anand AK, Punnakal AU, Chaudhoory AR, Garg C, Munjal RK, Bansal AK, et al. Stereotactic body radiation therapy (SBRT) for lung and liver tumours. *J Indian Med Assoc* 2012;**110**:462–4.
254. Barabino M, Santambrogio R, Giovenzana M, Polizzi M, Torri F, Opocher E. Clinical outcomes of laparoscopic radio-frequency or microwave ablation for hepatocellular carcinoma. *Surg Endosc* 2018;**32**:S466. <http://dx.doi.org/10.1007/s00464-018-6180-6>
255. Bassanello M, Cillo U, Vitale A, Lumachi F, Ciarleglio FA, Boccagni P, et al. Multimodal approach and its impact on survival for patients with hepatocellular carcinoma. *Anticancer Res* 2003;**23**:4047–53.
256. Beyer LP, Lurken L, Verloh N, Haimerl M, Michalik K, Schaible J, et al. Stereotactically navigated percutaneous microwave ablation (MWA) compared to conventional MWA: a matched pair analysis. *Int J Comput Assist Radiol Surg* 2018;**13**:1991–7. <https://dx.doi.org/10.1007/s11548-018-1778-7>
257. Bhutiani N, Philips P, Scoggins CR, McMasters KM, Potts MH, Martin RCG. Evaluation of tolerability and efficacy of irreversible electroporation (IRE) in treatment of Child–Pugh B (7/8) hepatocellular carcinoma (HCC). *HPB (Oxford)* 2016;**18**:593–9. <https://dx.doi.org/10.1016/j.hpb.2016.03.609>
258. Borzio M, Colloredo G, Pioltelli P, Quagliuolo M, Gruppo E, Lombardo. Epidemiology and outcome of hepatocellular carcinoma in Lombardy. *Dig Liver Dis* 2007;**39**:1011–7.
259. Bouda D, Barrau V, Raynaud L, Dioguardi Burgio M, Paulatto L, Roche V, et al. Factors associated with tumor progression after percutaneous ablation of hepatocellular carcinoma: comparison between monopolar radiofrequency and microwaves. Results of a propensity score matching analysis. *Cardiovasc Intervent Radiol* 2020;**43**:1608–18. <https://dx.doi.org/10.1007/s00270-020-02549-8>
260. Bu X, Ge Z, Ma J, Guo S, Wang Y, Liu J. Long-term efficacy of radiofrequency ablation compared to surgical resection for the treatment of small hepatocellular carcinoma. *J BUON* 2015;**20**:548–54.
261. Bujold A, Massey C, Kim JJ, Brierley JD, Cho C, Wong R, et al. Outcomes of stereotactic body radiotherapy (SBRT) for hepatocellular carcinoma (HCC). *Int J Radiat Oncol Biol Phys* 2011;**81**:S70–S1.
262. Casaccia M, Santori G, Bottino G, Diviacco P, Andorno E. Laparoscopic resection vs laparoscopic radiofrequency ablation for the treatment of small hepatocellular carcinomas: a single-center analysis. *World J Gastroenterol* 2017;**23**:653–60. <https://dx.doi.org/10.3748/wjg.v23.i4.653>
263. Chagnon S. Prospective randomized trial of radiofrequency ablation versus surgical resection for small hepatocellular carcinomas. *J Radiol* 2007;**88**:1842–4. <http://dx.doi.org/10.1016/S0221-0363%2807%2978361-0>

264. Chen KY, Xiang GA, Wang HN, Xiao FL. Intraoperative radiofrequency ablation and 125I therapy for preventing local recurrence in hepatocellular carcinoma after hepatectomy. *Zhonghua Zhong Liu Za Zhi* 2007;**29**:626–8.
265. Chen LC, Chiou WY, Lin HY, Lee MS, Lo YC, Huang LW, *et al.* Comparing stereotactic ablative radiotherapy (SABR) versus re-trans-catheter arterial chemoembolization (re-TACE) for hepatocellular carcinoma patients who had incomplete response after initial TACE (TASABR): a randomized controlled trial. *BMC Cancer* 2019;**19**:275. <http://dx.doi.org/10.1186/s12885-019-5461-3>
266. Cheung TT, Tp PR, Tat F, Lo CM. Survival analysis of high intensity focused ultrasound ablation in patients with small HCC. *HPB (Oxford)* 2012;**14**:42. <http://dx.doi.org/10.1111/j.1477-2574.2012.00510.x>
267. Cheung TT, Fan ST, Chan SC, Chok KS, Chu FS, Jenkins CR, *et al.* High-intensity focused ultrasound ablation: an effective bridging therapy for hepatocellular carcinoma patients. *World J Gastroenterol* 2013;**19**:3083–9. <https://dx.doi.org/10.3748/wjg.v19.i20.3083>
268. Cho HJ, Kim JK, Nam JS, Wang HJ, Lee JH, Kim BW, *et al.* High circulating microRNA-122 expression is a poor prognostic marker in patients with hepatitis B virus-related hepatocellular carcinoma who undergo radiofrequency ablation. *Clin Biochem* 2015;**48**:1073–8. <http://dx.doi.org/10.1016/j.clinbiochem.2015.06.019>
269. Chong CC, Lee KF, Chu CM, Chan AW, Yu SC, Lai PB. Laparoscopic hepatectomy (with or without robotic assistance) versus radiofrequency ablation as a minimally invasive treatment for very early-stage or early-stage hepatocellular carcinoma. *Dig Surg* 2020;**37**:65–71. <https://dx.doi.org/10.1159/000497112>
270. Cillo U, Noaro G, Vitale A, Neri D, D'Amico F, Gringeri E, *et al.* Laparoscopic microwave ablation in patients with hepatocellular carcinoma: a prospective cohort study. *HPB (Oxford)* 2014;**16**:979–86. <https://dx.doi.org/10.1111/hpb.12264>
271. Costa M, Santambrogio R, Barabino M, Zuin M, Bertolini E, Opocher E. Microinvasive intraoperative ultrasound pattern in predicting outcomes for single small (<3 cm) hepatocellular carcinoma. *J Hepatol* 2015;**62**:S319–20.
272. de Geus SW, Kasumova G, Kent TS, Ng SC, McAneny D, Sachs T, *et al.* Ablation versus hepatectomy for early-stage hepatocellular carcinoma: a matched nationwide review. *Gastroenterology* 2018;**154**:S-1279. <http://dx.doi.org/10.1016/S0016-5085%2818%2934201-X>
273. Denecke T, Stelter L, Schnapauff D, Steffen I, Sinn B, Schott E, *et al.* CT-guided interstitial brachytherapy of hepatocellular carcinoma before liver transplantation: an equivalent alternative to transarterial chemoembolization? *Eur Radiol* 2015;**25**:2608–16. <http://dx.doi.org/10.1007/s00330-015-3660-0>
274. Di Sandro S, Benuzzi L, Lauterio A, Botta F, De Carlis R, Najjar M, *et al.* Single hepatocellular carcinoma approached by curative-intent treatment: a propensity score analysis comparing radiofrequency ablation and liver resection. *Eur J Surg Oncol* 2019;**45**:1691–9. <https://dx.doi.org/10.1016/j.ejso.2019.04.023>
275. Ding W, Yu J, Liu F, Yu X, Cheng Z, Han Z, *et al.* Percutaneous microwave ablation versus robot-assisted hepatectomy for early hepatocellular carcinoma: a real-world single-center study. *Dig Liver Dis* 2021;**06**:06. <https://dx.doi.org/10.1016/j.dld.2021.04.008>
276. Ei S, Tanabe M, Hibi T, Itano O, Shinoda M, Kitago M, *et al.* Cryoablation compared with radiofrequency ablation and microwave coagulation therapy for the local control of primary hepatocellular carcinoma. *J Clin Oncol* 2013;**31**:286.

277. Eloubeidi MA, Provenzale D. CT-guided treatment of ultrasonically invisible hepatocellular carcinoma. *Am J Gastroenterol* 2000;**95**:2102–6. <http://dx.doi.org/10.1016/S0002-9270%2800%2901067-4>
278. Freeman E, Cheung W, Ferdousi S, Kavnoudias H, Majeed A, Kemp W, *et al.* Irreversible electroporation versus radiofrequency ablation for hepatocellular carcinoma: a single centre propensity-matched comparison. *Scand J Gastroenterol* 2021;**56**:942–7. <https://dx.doi.org/10.1080/00365521.2021.1930145>
279. Gannon CJ, Izzo F, Aloia TA, Pignata S, Nasti G, Vallone P, *et al.* Can hepatocellular cancer screening increase the proportion of long-term survivors? *Hepatogastroenterology* 2009;**56**:1152–6.
280. Garavoglia M, Oldani A, Gentili S, Portigliotti L, D'Agostino G. Percutaneous radiofrequency ablation versus surgical radiofrequency-assisted nodulectomy in treatment of small single nodes of hepatocellular carcinoma: our experience. *Minerva Chir* 2013;**68**:367–75.
281. Ghweil A, Osman H. Percutaneous microwave and radiofrequency ablation for hepatocellular carcinoma: a comparative study. *Hepatol Int* 2019;**13**:S161–2. <http://dx.doi.org/10.1007/s12072-019-09936-5>
282. Guibal A, Bertin C, Egels S, Savier E, Grenier PA, Lucidarme O. Contrast-enhanced ultrasound (CEUS) follow-up after radiofrequency ablation or cryoablation of focal liver lesions: treated-area patterns and their changes over time. *Eur Radiol* 2013;**23**:1392–400. <https://dx.doi.org/10.1007/s00330-012-2702-0>
283. Guo WX, Zhai B, Lai EC, Li N, Shi J, Lau WY, *et al.* Percutaneous radiofrequency ablation versus partial hepatectomy for multicentric small hepatocellular carcinomas: a nonrandomized comparative study. *World J Surg* 2010;**34**:2671–6. <https://dx.doi.org/10.1007/s00268-010-0732-9>
284. Hara K, Takeda A, Tsurugai Y, Sanuki N, Saigusa Y, Maeda S, *et al.* Clinical outcomes after treatment for hepatocellular carcinoma, stereotactic body radiotherapy vs radiofrequency ablation: a propensity score-matched analysis. *Hepatology* 2018;**68**:848A. <http://dx.doi.org/10.1002/hep.30257>
285. Hara K, Takeda A, Tsurugai Y, Saigusa Y, Sanuki N, Eriguchi T, *et al.* Radiotherapy for hepatocellular carcinoma results in comparable survival to radiofrequency ablation: a propensity score analysis. *Hepatology* 2019;**69**:2533–45. <https://dx.doi.org/10.1002/hep.30591>
286. Hasegawa K, Kokudo N, Makuuchi M, Izumi N, Ichida T, Kudo M, *et al.* Comparison of resection and ablation for hepatocellular carcinoma: a cohort study based on a Japanese nationwide survey. *J Hepatol* 2013;**58**:724–9. <https://dx.doi.org/10.1016/j.jhep.2012.11.009>
287. He Q, Jiang JJ, Jiang YX, Wang WT, Yang L, Liver S, *et al.* Health-related quality of life comparisons after radical therapy for early-stage hepatocellular carcinoma. *Transplant Proc* 2018;**50**:1470–4. <https://dx.doi.org/10.1016/j.transproceed.2018.04.041>
288. Helmberger T, Dogan S, Straub G, Schrader A, Jungst C, Reiser M, *et al.* Liver resection or combined chemoembolization and radiofrequency ablation improve survival in patients with hepatocellular carcinoma. *Digestion* 2007;**75**:104–12.
289. Hiraoka A, Horiike N, Yamashita Y, Koizumi Y, Doi K, Yamamoto Y, *et al.* Efficacy of radiofrequency ablation therapy compared to surgical resection in 164 patients in Japan with single hepatocellular carcinoma smaller than 3 cm, along with report of complications. *Hepatogastroenterology* 2008;**55**:2171–4.
290. Hiraoka A, Kumada T, Michitaka K, Toyoda H, Tada T, Takaguchi K, *et al.* Clinical features of hemodialysis patients treated for hepatocellular carcinoma: comparison between resection and radiofrequency ablation. *Mol Clin Oncol* 2017;**6**:455–61. <https://dx.doi.org/10.3892/mco.2017.1192>

291. Ho EY, Lerrigo R, Arunkumar A, Shen H, Trimble EL, Corvera CU, *et al.* Hepatocellular carcinoma with a plan for surgical resection is associated with improved survival compared to a plan for liver transplantation. *Gastroenterology* 2012;**142**:S985.
292. Hsu CY, Lee YH, Hsia CY, Huang YH, Su CW, Lin HC, *et al.* Performance status enhances the selection of treatment for patients with hepatocellular carcinoma within the Milan criteria. *Ann Surg Oncol* 2013;**20**:2035–42. <https://dx.doi.org/10.1245/s10434-012-2847-8>
293. Huang WY, Jen YM, Lee MS, Chang LP, Chen CM, Ko KH, *et al.* Stereotactic body radiation therapy in recurrent hepatocellular carcinoma. *Int J Radiat Oncol Biol Phys* 2012;**84**:355–61. <https://dx.doi.org/10.1016/j.ijrobp.2011.11.058>
294. Ikeda K, Kobayashi M, Saitoh S, Someya T, Hosaka T, Sezaki H, *et al.* Cost-effectiveness of radiofrequency ablation and surgical therapy for small hepatocellular carcinoma of 3cm or less in diameter. *Hepatol Res* 2005;**33**:241–9.
295. Ikeda K, Seki T, Umehara H, Inokuchi R, Tamai T, Sakaida N, *et al.* Clinicopathologic study of small hepatocellular carcinoma with microscopic satellite nodules to determine the extent of tumor ablation by local therapy. *Int J Oncol* 2007;**31**:485–91.
296. Imai K, Beppu T, Chikamoto A, Doi K, Okabe H, Nitta H, *et al.* Radiofrequency ablation versus hepatic resection as an initial treatment for solitary small-sized hepatocellular carcinoma. *J Clin Oncol* 2012;**30**:290.
297. Ito T, Tanaka S, Iwai S, Takemura S, Agihara A, Uchida S, *et al.* Local control of laparoscopic hepatic resection for surface hepatocellular carcinoma: comparison with percutaneous radiofrequency ablation. *Liver Cancer* 2015;**4**:169–70. <http://dx.doi.org/10.1159/000367745>
298. Ito T, Tanaka S, Iwai S, Takemura S, Hagihara A, Uchida-Kobayashi S, *et al.* Outcomes of laparoscopic hepatic resection versus percutaneous radiofrequency ablation for hepatocellular carcinoma located at the liver surface: a case-control study with propensity score matching. *Hepatol Res* 2016;**46**:565–74. <http://dx.doi.org/10.1111/hepr.12592>
299. Jianyong L, Jinjing Z, Lunan Y, Jingqiang Z, Wentao W, Yong Z, *et al.* Preoperative adjuvant transarterial chemoembolization cannot improve the long term outcome of radical therapies for hepatocellular carcinoma. *Sci Rep* 2017;**7**:41624. <https://dx.doi.org/10.1038/srep41624>
300. Juloori A, Liao CY, Lemons JM, Singh AK, Iyer R, Robbins JR, *et al.* Phase I study of stereotactic body radiotherapy followed by ipilimumab with nivolumab vs. nivolumab alone in unresectable hepatocellular carcinoma. *Int J Radiat Oncol Biol Phys* 2020;**108**:S149–50. <https://dx.doi.org/10.1016/j.ijrobp.2020.07.900>
301. Kang TW, Rhim H. Radiofrequency ablation versus non-anatomical resection: long-term results in 580 patients with a single small hepatocellular carcinoma of  $\leq 3$ cm. *Cardiovasc Intervent Radiol* 2014;**37**:S253.
302. Kanwal F, Befeler A, Chari RS, Marrero J, Kahn J, Afdhal N, *et al.* Potentially curative treatment in patients with hepatocellular cancer – results from the Liver Cancer Research Network. *Aliment Pharmacol Ther* 2012;**36**:257–65. <https://dx.doi.org/10.1111/j.1365-2036.2012.05174.x>
303. Katsoulakis E, Riaz N, Cannon DM, Goodman K, Spratt DE, Lovelock M, *et al.* Image-guided radiation therapy for liver tumors: gastrointestinal histology matters. *Am J Clin Oncol* 2014;**37**:561–7. <https://dx.doi.org/10.1097/COC.0b013e318282a86b>
304. Kawamura Y, Ikeda K, Shindoh J, Kobayashi Y, Kasuya K, Fujiyama S, *et al.* No-touch ablation in hepatocellular carcinoma has the potential to prevent intrasubsegmental recurrence to the same degree as surgical resection. *Hepatol Res* 2019;**49**:164–76. <https://dx.doi.org/10.1111/hepr.13254>

305. Kawaoka T, Aikata H, Morio K, Nakahara T, Murakami E, Tsuge M, *et al.* Stereotactic body radiation therapy combined with transcatheter arterial chemoembolization for small hepatocellular carcinoma. *Hepatology* 2019;**70**:528A.
306. Kennedy JE, Wu F, ter Haar GR, Gleeson FV, Phillips RR, Middleton MR, *et al.* High-intensity focused ultrasound for the treatment of liver tumours. *Ultrasonics* 2004;**42**:931–5. <https://dx.doi.org/10.1016/j.ultras.2004.01.089>
307. Kim GA, Shim JH, Kim SY, Won HJ, Shin YM, Kim PN, *et al.* Surgical resection versus radiofrequency ablation in the treatment of patients with solitary hepatocellular carcinoma without portal hypertension: a 10-year outcome study. *J Hepatol* 2014;**60**:S53. <http://dx.doi.org/10.1016/S0168-8278%2814%2960130-6>
308. Kim JM, Kang TW, Choi JY, Lee S, Cho W, Kwon CHD, *et al.* Single HCC  $\leq$ 3cm in left lateral segment: liver resection or radiofrequency ablation? *J Hepatol* 2014;**60**:S396. [https://doi.org/10.1016/S0168-8278\(14\)61123-5](https://doi.org/10.1016/S0168-8278(14)61123-5)
309. Kim DS, Jung SW, Yu YD, Han JH, Suh SO. Intention-to-treat analysis of various treatment outcomes for hepatocellular carcinoma defined BCLC stage a with portal hypertension. *HPB (Oxford)* 2016;**18**:e218.
310. Kim HJ, Park S, Kim KJ, Seong J. Clinical significance of soluble programmed cell death ligand-1 (sPD-L1) in hepatocellular carcinoma patients treated with radiotherapy. *Radiother Oncol* 2018;**129**:130–5. <https://dx.doi.org/10.1016/j.radonc.2017.11.027>
311. Kim TH, Park JW, Kim BH, Kim H, Moon SH, Kim SS, *et al.* Does risk-adapted proton beam therapy have a role as a complementary or alternative therapeutic option for hepatocellular carcinoma? *Cancers* 2019;**11**:230. <https://dx.doi.org/10.3390/cancers11020230>
312. Kimura T, Aikata H, Doi Y, Imano N, Takeuchi Y, Takahashi I, *et al.* Comparison of stereotactic body radiation therapy combined with or without transcatheter arterial chemoembolization for patients with small hepatocellular carcinoma ineligible for resection or ablation therapies. *Technol Cancer Res Treat* 2018;**17**:1–11. <https://dx.doi.org/10.1177/1533033818783450>
313. Kimura T, Doi Y, Aikata H, Imano N, Takeuchi Y, Takahashi I, *et al.* Comparison of stereotactic body radiation therapy combined with or without transcatheter arterial chemoembolization for patients with small hepatocellular carcinoma ineligible for resection or ablation therapies. *Int J Radiat Oncol Biol Phys* 2018;**102**:E57. <https://dx.doi.org/10.1016/j.ijrobp.2018.07.481>
314. Ko SE, Lee MW, Rhim H, Kang TW, Song KD, Cha DI, *et al.* Comparison of procedure-related complications between percutaneous cryoablation and radiofrequency ablation for treating periductal hepatocellular carcinoma. *Int J Hyperthermia* 2020;**37**:1354–61. <https://dx.doi.org/10.1080/02656736.2020.1849824>
315. Komatsu S, Fukumoto T, Demizu Y, Miyawaki D, Terashima K, Sasaki R, *et al.* Clinical results and risk factors of proton and carbon ion therapy for hepatocellular carcinoma. *Cancer* 2011;**117**:4890–904. <https://dx.doi.org/10.1002/cncr.26134>
316. Komatsu S, Terashima K, Matsuo Y, Takahashi D, Suga M, Nishimura N, *et al.* Validation of combination treatment with surgical spacer placement and subsequent particle radiotherapy for unresectable hepatocellular carcinoma. *J Surg Oncol* 2019;**120**:214–22. <https://dx.doi.org/10.1002/jso.25495>
317. Kooby DA, Egnatashvili V, Srinivasan S, Chamsuddin A, Delman KA, Kauh J, *et al.* Comparison of yttrium-90 radioembolization and transcatheter arterial chemoembolization for the treatment of unresectable hepatocellular carcinoma. *J Vasc Interv Radiol* 2010;**21**:224–30. <http://dx.doi.org/10.1016/j.jvir.2009.10.013>

318. Kuang M, Xie XY, Huang C, Wang Y, Lin MX, Xu ZF, *et al.* Long-term outcome of percutaneous ablation in very early-stage hepatocellular carcinoma. *J Gastrointest Surg* 2011;**15**:2165–71. <https://dx.doi.org/10.1007/s11605-011-1716-2>
319. Kudithipudi V, Day E, Thai N, Kirichenko A. Liver stereotactic radiotherapy (SRT) with functional treatment planning for patients with intermediate stage hepatocellular carcinoma (HCC). *J Radiat Oncol* 2017;**6**:371–7. <http://dx.doi.org/10.1007/s13566-017-0325-4>
320. Kuo YH, Lu SN, Chen CL, Cheng YF, Lin CY, Hung CH, *et al.* Hepatocellular carcinoma surveillance and appropriate treatment options improve survival for patients with liver cirrhosis. *Eur J Cancer* 2010;**46**:744–51. <http://dx.doi.org/10.1016/j.ejca.2009.12.018>
321. Kuo MJ, Mo LR, Chen CL. Factors predicting long-term outcomes of early-stage hepatocellular carcinoma after primary curative treatment: the role of surgical or nonsurgical methods. *BMC Cancer* 2021;**21**:250. <https://dx.doi.org/10.1186/s12885-021-07948-9>
322. Kuromatsu R, Takata A, Fukushima N, Sumie S, Nakano M. DCP is an important risk factor for recurrence after radiofrequency ablation of single hepatocellular carcinoma < 3cm in diameter. *Hepatol Int* 2009;**3**:151. <http://dx.doi.org/10.1007/s12072-009-9123-4>
323. Kwon JH, Lee JH, Cho EJ, Lee YB, Cho Y, Lee DH, *et al.* The efficacy of transarterial chemoembolization for small hepatic nodules with atypical enhancement on contrast-enhanced ct but typical enhancement on gadoxetic acid-enhanced MRI in chronic liver disease patients: a post hoc analysis of prospective study. *Hepatology* 2012;**56**:482A. <http://dx.doi.org/10.1002/hep.26040>
324. Lapinski TW, Tarasik A, Januszkiewicz M, Flisiak R. Clinical aspects and treatment of hepatocellular carcinoma in north-eastern Poland. *Clin Exp Hepatol* 2021;**7**:79–84. <http://dx.doi.org/10.5114/ceh.2021.104631>
325. Lee G, Choi MS, Song L, Min YW, Gwak GY, Paik YH, *et al.* Optimal treatment policy for elderly patients with single small hepatocellular carcinoma: comparison with young patients after propensity score matching. *J Gastroenterol Hepatol* 2012;**27**:266. <http://dx.doi.org/10.1111/jgh.12006>
326. Lee G, Choi MS, Sinn DH, Gwak GY, Paik YH, Lee JH, *et al.* Long-term outcomes after treatment of single small hepatocellular carcinoma in elderly patients with well-preserved liver function and good performance status. *United Eur Gastroenterol J* 2014;**2**:A149. <http://dx.doi.org/10.1177/2050640614548980>
327. Lee HA, Lee YR, Kim T, Young YS, Lee YS, Suh SJ, *et al.* Surgical resection versus radiofrequency ablation in single small hepatocellular carcinoma: data from Korea Central Cancer Registry for Hepatocellular Carcinoma database. *Hepatology* 2018;**68**:841A–2A. <http://dx.doi.org/10.1002/hep.30257>
328. Lee DH, Lee JM, Kim JW, Lee MW, Kim JM. Laparoscopic liver resection vs. percutaneous radiofrequency ablation for small single nodular HCC: comparison of treatment outcome. *J Hepatol* 2019;**70**:611–3. <http://dx.doi.org/10.1016/S0168-8278%2819%2930204-1>
329. Lee DH, Kim JW, Lee JM, Kim JM, Lee MW, Rhim H, *et al.* Laparoscopic liver resection versus percutaneous radiofrequency ablation for small single nodular hepatocellular carcinoma: comparison of treatment outcomes. *Liver Cancer* 2021;**10**:25–37. <https://dx.doi.org/10.1159/000510909>
330. Lei JY, Wang WT, Yan LN, Wen TF, Li B. Radiofrequency ablation versus surgical resection for small unifocal hepatocellular carcinomas. *Medicine (Baltim)* 2014;**93**:e271. <https://dx.doi.org/10.1097/MD.0000000000000271>

331. Li S, Wang Y, Jun D, Wu P. Comparison of the effectiveness of percutaneous microwave ablation versus hepatectomy for hepatocellular carcinoma. *Zhonghua Zhong Liu Za Zhi* 2015;**37**:301–7. <http://dx.doi.org/10.3760/cma.j.issn.0253-3766.2015.04.013>
332. Li XL, Guo WX, Hong XD, Yang L, Wang K, Shi J, *et al.* Efficacy of the treatment of transarterial chemoembolization combined with radiotherapy for hepatocellular carcinoma with portal vein tumor thrombus: a propensity score analysis. *Hepatol Res* 2016;**46**:1088–98. <https://dx.doi.org/10.1111/hepr.12657>
333. Li IF, Huang JC, Chen JJ, Wang TE, Huang SS, Tsay SL. Factors related to the quality of life in liver cancer patients during treatment phase: a follow-up study. *Eur J Cancer Care* 2019;**28**:e13146. <https://dx.doi.org/10.1111/ecc.13146>
334. Liu PH, Hsu CY, Hsia CY, Lee YH, Huang YH, Chiou YY, *et al.* Surgical resection versus radiofrequency ablation for single hepatocellular carcinoma  $\leq$  2 cm in a propensity score model. *Ann Surg* 2016;**263**:538–45. <https://dx.doi.org/10.1097/SLA.0000000000001178>
335. Liu PH, Huo TI. Treating very early-stage HCC: have we found the holy grail? *Ann Surg* 2017;**266**:e31–2. <http://dx.doi.org/10.1097/SLA.0000000000001349>
336. Liu W, Zheng Y, He W, Zou R, Qiu J, Shen J, *et al.* Microwave vs radiofrequency ablation for hepatocellular carcinoma within the Milan criteria: a propensity score analysis. *Aliment Pharmacol Ther* 2018;**48**:671–81. <https://dx.doi.org/10.1111/apt.14929>
337. Liu W, Li B, Yuan Y. Letter: is microwave ablation superior to radiofrequency ablation for early stage hepatocellular carcinoma? Authors' reply. *Aliment Pharmacol Ther* 2018;**48**:1326–7. <https://dx.doi.org/10.1111/apt.15048>
338. Liu H, Lee D, McLean K, Leggett D, Hodgkinson P, Fawcett J, *et al.* Australian experience of SBRT in early and advanced stage hepatocellular carcinoma. *Radiother Oncol* 2019;**141**:S19–20.
339. Loo KF, Woodman R, Bogatic D, Muller K, Chandran V, Chinnaratha MA, *et al.* Treatment stage migration for early hepatocellular carcinoma and its impact on outcomes: a multicenter study. *JGH Open* 2020;**35**:67–8. <https://dx.doi.org/10.1111/jgh.15269>
340. Lu SN, Wang JH, Chen CL. Survival comparison between surgical resection and radiofrequency ablation in hepatocellular carcinoma patients with BCLC (very) early stage. *J Hepatol* 2010;**52**:S94–S5. <http://dx.doi.org/10.1016/S0168-8278%2810%2960225-5>
341. Maezawa K, Midorikawa T, Kigawa G, Ishibashi K, Hatakeyama T, Miyakawa K, *et al.* A role of laparoscopic microwave coagulation therapy for hepatocellular carcinoma as a less invasive surgery. *Nihon Shokaki Geka Gakkai Zasshi* 2005;**38**:279–88. <http://dx.doi.org/10.5833/jjgs.38.279>
342. Mariani NM, Santambrogio R, Barabino M, Scifo G, Giovenzana M, Opocher E. Laparoscopic liver resection vs radiofrequency ablation in cirrhotic patients with hepatocellular carcinoma: towards the 'gold standard'. *Surg Endosc* 2017;**31**:S33. <http://dx.doi.org/10.1007/s00464-017-5540-y>
343. Martin AN, Das D, Johnston LE, Bauer TW, Adams RB, Zaydfudim VM. Management strategies for patients with solitary hepatocellular carcinoma  $\leq$  3cm. *Ann Surg Oncol* 2016;**23**:S104–5. <http://dx.doi.org/10.1245/s10434-015-5010-5>
344. Merle P, Beziat C, Wautot V, Trepo C, Mornex F. Tolerance and efficacy of high-dose 3D-conformal radiation therapy (CRT) in cirrhotic patients with small hepatocellular carcinomas (HCC) not suitable for curative therapies. *Hepatology* 2005;**42**:394A–5A.
345. Ming W. Survival comparison between surgical resection and percutaneous radiofrequency ablation for patients in BCLC early stage hepatocellular carcinoma. *Hepatol Int* 2012;**6**:220. <http://dx.doi.org/10.1007/s12072-011-9333-4>

346. Mizumoto M, Okumura T, Hashimoto T, Fukuda K, Oshiro Y, Fukumitsu N, *et al.* Proton beam therapy for hepatocellular carcinoma: a comparison of three treatment protocols. *Int J Radiat Oncol Biol Phys* 2011;**81**:1039–45. <https://dx.doi.org/10.1016/j.ijrobp.2010.07.015>
347. Mo QG, Liang AM, Yang NW, Zhao YN, Yuan WP. Surgery-predominant comprehensive therapy for 134 patients with small hepatocellular carcinoma. *Ai Zheng* 2003;**22**:189–91.
348. Mohnike K, Steffen IG, Seidensticker M, Hass P, Damm R, Peters N, *et al.* Radioablation by image-guided (HDR) brachytherapy and transarterial chemoembolization in hepatocellular carcinoma: a randomized phase II trial. *Cardiovasc Intervent Radiol* 2019;**42**:239–49. <http://dx.doi.org/10.1007/s00270-018-2127-5>
349. Molinari M, Walsh MJ, Renfrew PD, Kneteman NM, Helton W. Treatment trade-off for ablation or hepatic resection for small HCC. *Hepatology* 2009;**50**:1098A. <http://dx.doi.org/10.1002/hep.23307>
350. Molinari M, De Coutere S, Krahn M, Helton S, Urbach DR. Patients' preferences and trade-offs for the treatment of early stage hepatocellular carcinoma. *J Surg Res* 2014;**189**:57–67. <https://dx.doi.org/10.1016/j.jss.2014.02.015>
351. Mornex F, Merle P, Beziat C, Wautot V, Kubas A, Khodri M, *et al.* Tolerance and efficacy of high-dose 3D-conformal radiation therapy (CRT) in cirrhotic patients with small hepatocellular carcinomas (HCC) not suitable for curative therapies. *Int J Radiat Oncol Biol Phys* 2005;**63**:S15. <https://dx.doi.org/10.1016/j.ijrobp.2005.07.032>
352. Nahon P, Seror O, Oberti F, Dimby SF, Aube C, Blanc JF, *et al.* Radiological and pathological response to neoadjuvant nivolumab in patients with BCLC A HCC treated by curative percutaneous irreversible electroporation: preliminary report from the NIVOLEP trial (NCT03630640). *J Hepatol* 2021;**75**:S198.
353. Nanashima A, Yamaguchi H, Omagari K, Nakazaki T, Aritomi T, Hatano K, *et al.* A comparison of hepatic resection and ablative therapy regarding the survival of hepatocellular carcinoma patients in Nagasaki. *Acta Med Nagasaki* 2004;**49**:87–91.
354. Nanashima A, Tobinaga S, Masuda J, Miyaaki H, Taura N, Takeshita H, *et al.* Selecting treatment for hepatocellular carcinoma based on the results of hepatic resection and local ablation therapy. *J Surg Oncol* 2010;**101**:481–5. <https://dx.doi.org/10.1002/jso.21523>
355. Nathan H, Segev DL, Bridges JF, Massie AB, Cameron AM, Hirose K, *et al.* Influence of nonclinical factors on choice of therapy for early hepatocellular carcinoma. *Ann Surg Oncol* 2013;**20**:448–56. <https://dx.doi.org/10.1245/s10434-012-2619-5>
356. Ogiso S, Seo S, Eso Y, Yoh T, Kawai T, Okumura S, *et al.* Laparoscopic liver resection versus percutaneous radiofrequency ablation for small hepatocellular carcinoma. *HPB (Oxford)* 2021;**23**:533–7. <https://dx.doi.org/10.1016/j.hpb.2020.08.009>
357. Oh JH, Sinn DH, Kang W, Gwak GY, Paik YH, Lee JH, *et al.* Long-term outcome according to initial treatment modality in multiple hepatocellular carcinoma within the Milan criteria. *J Hepatol* 2019;**70**:615. <http://dx.doi.org/10.1016/S0168-8278%2819%2930204-1>
358. Oh JH, Sinn DH, Choi GS, Kim JM, Joh JW, Kang TW, *et al.* Comparison of outcome between liver resection, radiofrequency ablation, and transarterial therapy for multiple small hepatocellular carcinoma within the Milan criteria. *Ann Surg Treat Res* 2020;**99**:238–46. <https://dx.doi.org/10.4174/astr.2020.99.4.238>
359. Ohmoto K, Yoshioka N, Tomiyama Y, Shibata N, Kawase T, Yoshida K, *et al.* Thermal ablation therapy for hepatocellular carcinoma: comparison between radiofrequency ablation and percutaneous microwave coagulation therapy. *Hepatogastroenterology* 2006;**53**:651–4.

360. Pan YX, Long Q, Yi MJ, Chen JB, Chen JC, Zhang YJ, *et al.* Radiofrequency ablation versus laparoscopic hepatectomy for hepatocellular carcinoma: a real world single center study. *Eur J Surg Oncol* 2020;**46**:548–59. <https://dx.doi.org/10.1016/j.ejso.2019.10.026>
361. Park KW, Park JW, Kim TH, Choi JI, Kim SH, Park HS, *et al.* Five-year survival analysis of a cohort of hepatocellular carcinoma patients who treated at the National Cancer Center, Korea. *Korean J Hepatol* 2007;**13**:530–42.
362. Park S, Lee EJ, Rim CH, Seong J. Plasma cell-free DNA as a predictive marker after radiotherapy for hepatocellular carcinoma. *Yonsei Med J* 2018;**59**:470–9. <https://dx.doi.org/10.3349/ymj.2018.59.4.470>
363. Park JW, Kim TH, Koh YH, Kim BH, Kim MJ, Lee JH, *et al.* Proton beam radiotherapy versus radiofrequency ablation treatment in patients with recurrent hepatocellular carcinoma: a randomized controlled phase 3 non-inferiority APROH trial. *J Hepatol* 2020;**73**:S119. [https://dx.doi.org/10.1016/s0168-8278\(20\)30756-x](https://dx.doi.org/10.1016/s0168-8278(20)30756-x)
364. Peng ZW, Zhang YJ, Chen MS. Radiofrequency ablation versus hepatic resection for the treatment of hepatocellular carcinoma small than 2 cm. *HPB (Oxford)* 2011;**13**:124. <http://dx.doi.org/10.1111/j.1477-2574.2011.00383.x>
365. Peng Z, Chen M. Radiofrequency ablation versus liver resection for elderly patients with hepatocellular carcinoma (HCC) within the Milan criteria. *J Vasc Interv Radiol* 2012;**23**:853. e3–4. <http://dx.doi.org/10.1016/j.jvir.2012.04.029>
366. Peng ZW, Liu FR, Ye S, Xu L, Zhang YJ, Liang HH, *et al.* Radiofrequency ablation versus open hepatic resection for elderly patients (> 65 years) with very early or early hepatocellular carcinoma. *Cancer* 2013;**119**:3812–20. <https://dx.doi.org/10.1002/cncr.28293>
367. Peng JR, Zhu WH, Xu QQ, Gan LY, Zheng SM, Zhang DF, *et al.* Comparison of the therapeutic effects of microwave ablation and surgical resection for small hepatocellular carcinoma. *J Clin Oncol* 2014;**32**:e15124.
368. Praktijnjo M, Krabbe V, Pohlmann A, Sampels M, Jansen C, Meyer C, *et al.* Evolution of nodule stiffness might predict response to local ablative therapy: a series of patients with hepatocellular carcinoma. *PLOS ONE* 2018;**13**:e0192897. <http://dx.doi.org/10.1371/journal.pone.0192897>
369. Pryor D, Liu H, Hodgkinson P, Leggett D, McLean K, Lee D, *et al.* Efficacy and toxicity of stereotactic radiation therapy in early and advanced stage hepatocellular carcinoma. *J Gastroenterol Hepatol* 2019;**34**:34–5.
370. Rong W, Yu W, Wang L, Wu F, Zhang K, Chen B, *et al.* Adjuvant radiotherapy in central hepatocellular carcinoma after narrow-margin hepatectomy: a 10-year real-world evidence. *Chin J Cancer Res* 2020;**32**:645–53. <http://dx.doi.org/10.21147/j.issn.1000-9604.2020.05.09>
371. Ruzzenente A, Guglielmi A, Sandri M, Campagnaro T, Valdegamberi A, Conci S, *et al.* Surgical resection versus local ablation for HCC on cirrhosis: results from a propensity case-matched study. *J Gastrointest Surg* 2012;**16**:301–11. <https://dx.doi.org/10.1007/s11605-011-1745-x>
372. Ryu T, Takami Y, Wada Y, Tateishi M, Hara K, Yoshitomi M, *et al.* Hepatic resection versus surgical microwave ablation for single hepatocellular carcinoma ≤ 5cm in a propensity score matching. *Hepatology* 2018;**68**:849A. <http://dx.doi.org/10.1002/hep.30257>
373. Ryu T, Takami Y, Wada Y, Hara T, Sasaki S, Saitsu H. Hepatic resection versus operative microwave ablation for single hepatocellular carcinoma ≤ 5 cm: a propensity score-matched analysis. *Surgery* 2019;**166**:254–62. <https://dx.doi.org/10.1016/j.surg.2019.05.007>
374. Sako K, Komorizono Y, Oketani M, Hasegawa S, Yamasaki N, Shibatou T, *et al.* Outcome of patients with hepatitis C virus-related single, small hepatocellular carcinoma. *Anticancer Res* 2003;**23**:4191–6.

375. Santambrogio R, Costa M, Conti M, Zuin M, Opocher E, Bruno S. The impact of radiofrequency ablation and liver resection on hepatocellular carcinoma (HCC) recurrence and survival in cirrhotic patients with early and very early HCC: a retrospective-prospective cohort study. *Dig Liver Dis* 2009;**41**:A8–9. <https://doi.org/10.1016/j.dld.2009.02.025>
376. Santambrogio R, Chiang J, Barabino M, Meloni FM, Bertolini E, Melchiorre F, *et al.* Comparison of laparoscopic microwave to radiofrequency ablation of small hepatocellular carcinoma ( $\leq 3$  cm). *Ann Surg Oncol* 2017;**24**:257–63. <https://dx.doi.org/10.1245/s10434-016-5527-2>
377. Santambrogio R, Barabino M, Bruno S, Mariani N, Maroni N, Bertolini E, *et al.* Surgical resection vs. ablative therapies through a laparoscopic approach for hepatocellular carcinoma: a comparative study. *J Gastrointest Surg* 2018;**22**:650–60. <https://dx.doi.org/10.1007/s11605-017-3648-y>
378. Santambrogio R, Barabino M, D'Alessandro V, Iacob G, Opocher E, Gemma M, *et al.* Micronvasive behaviour of single small hepatocellular carcinoma: which treatment? *Updates Surg* 2021;**73**:1359–69. <http://dx.doi.org/10.1007/s13304-021-01036-0>
379. Schaible J, Lurken L, Wiggermann P, Verloh N, Einspieler I, Zeman F, *et al.* Primary efficacy of percutaneous microwave ablation of malignant liver tumors: comparison of stereotactic and conventional manual guidance. *Sci Rep* 2020;**10**:18835. <https://dx.doi.org/10.1038/s41598-020-75925-6>
380. Sheng L, Wang Y, Jun D, Peihong W. Comparison of the effectiveness of percutaneous microwave ablation versus hepatectomy for hepatocellular carcinoma. *Zhonghua Zhong Liu Za Zhi* 2015;**37**:301–7.
381. Shibata T, Maetani Y, Ametani F, Kubo T, Itoh K, Konishi J. Tumor ablation therapy for small hepatocellular carcinoma: comparison of radiofrequency ablation and percutaneous microwave coagulation therapy. *Radiology* 2001;**221**:248–9.
382. Shiozawa K, Watanabe M, Ikehara T, Matsukiyo Y, Kogame M, Kishimoto Y, *et al.* Comparison of percutaneous radiofrequency ablation and CyberKnife for initial solitary hepatocellular carcinoma: a pilot study. *World J Gastroenterol* 2015;**21**:13490–9. <http://dx.doi.org/10.3748/wjg.v21.i48.13490>
383. Simo KA, Sereika SE, Newton KN, Gerber DA. Laparoscopic-assisted microwave ablation for hepatocellular carcinoma: safety and efficacy in comparison with radiofrequency ablation. *J Surg Oncol* 2011;**104**:822–9. <https://doi.org/10.1002/jso.21933>
384. Song KD, Lim HK, Rhim H, Lee MW, Kang TW, Paik YH, *et al.* Hepatic resection vs percutaneous radiofrequency ablation of hepatocellular carcinoma abutting right diaphragm. *World J Gastrointest Oncol* 2019;**11**:227–37. <http://dx.doi.org/10.4251/wjgo.v11.i3.227>
385. Spangenberg HC. Resection or ablation in hepatocellular carcinoma: pro ablation. *Dtsch Med Wochenschr* 2008;**133**:1578. <http://dx.doi.org/10.1055/s-2008-1081113>
386. Stuart KA, Liu H, McLean K, Pryor D. Early Australian experience with stereotactic body radiation therapy in the management of hepatocellular carcinoma: efficacy and toxicity. *J Gastroenterol Hepatol* 2018;**33**:23–4.
387. Su TS, Liang P, Shixiong L, Zhou Y, Huang Y, Cheng T. Stereotactic body radiation therapy as a radical or salvage treatment for early-stage hepatocellular carcinoma: a prospective phase II study. *Int J Radiat Oncol Biol Phys* 2020;**108**:E582.
388. Suh SW, Lee KW, Lee JM, Yoo T, Choi YR, Yi NJ, *et al.* Prediction of aggressiveness in early-stage hepatocellular carcinoma for selection of surgical treatment. *Hepatology* 2013;**58**:1238A. <http://dx.doi.org/10.1002/hep.26883>

389. Sutter O, Fihri A, Ourabia-Belkacem R, Sellier N, Diallo A, Seror O. Real-time 3D virtual target fluoroscopic display for challenging hepatocellular carcinoma ablations using cone beam CT. *Technol Cancer Res Treat* 2018;**17**:1–7. <https://dx.doi.org/10.1177/1533033818789634>
390. Takamatsu S, Yamamoto K, Kawamura M, Asahi S, Satoh Y, Tameshige Y, *et al.* Early experience with proton beam therapy for hepatocellular carcinoma. *Int J Radiat Oncol Biol Phys* 2014;**90**:S370–1.
391. Takami Y, Wada Y, Saitsu H. Strategy for treatment of hepatocellular carcinoma - Hr and MCN. *Hepatol Int* 2009;**3**:79. <http://dx.doi.org/10.1007/s12072-008-9120-z>
392. Takami Y, Ryu T, Wada Y, Saitsu H. Evaluation of microwave coagulo-necrotic therapy (MCN) for hepatocellular carcinoma (HCC): a single center experience of 650 consecutive cases. *Hepatology* 2010;**52**:1153A. <http://dx.doi.org/10.1002/hep.23996>
393. Takayama T, Makuuchi M, Hasegawa K. Single HCC smaller than 2 cm: surgery or ablation? *J Hepatobiliary Pancreat Sci* 2010;**17**:422–4. <https://dx.doi.org/10.1007/s00534-009-0239-7>
394. Takayasu K, Arai S, Sakamoto M, Matsuyama Y, Kudo M, Kaneko S, *et al.* Impact of resection and ablation for single hypovascular hepatocellular carcinoma  $\leq$  2 cm analysed with propensity score weighting. *Liver Int* 2018;**38**:484–93. <https://dx.doi.org/10.1111/liv.13670>
395. Takeda A, Takahashi M, Kunieda E, Takeda T, Sanuki N, Koike Y, *et al.* Hypofractionated stereotactic radiotherapy with and without transarterial chemoembolization for small hepatocellular carcinoma not eligible for other ablation therapies: preliminary results for efficacy and toxicity. *Hepatol Res* 2008;**38**:60–9.
396. Takeda A, Sanuki N, Tsurugai Y, Iwabuchi S, Matsunaga K, Ebinuma H, *et al.* Phase 2 study of stereotactic body radiotherapy and optional transarterial chemoembolization for solitary hepatocellular carcinoma not amenable to resection and radiofrequency ablation. *Cancer* 2016;**122**:2041–9. <https://dx.doi.org/10.1002/cncr.30008>
397. Tanaka S, Iwai S, Takemura S, Hagiwara A, Uchida S, Ito T, *et al.* Laparoscopic hepatic resection versus radiofrequency ablation for local control of small hepatocellular carcinoma located on the surface of the liver. *HPB (Oxford)* 2015;**17**:37.
398. Tanguturi S, Kambadakone A, Niemierko A, Nguyen KN, Wo JY, Blaszkowsky LS, *et al.* Early response assessment after hypofractionated proton radiation therapy for primary liver tumors. *Int J Radiat Oncol Biol Phys* 2015;**93**:E125.
399. Tatineni T, Kataki K, Madhavan R, Das R, Dutta D. Prognostication of HCC with PVT treated with SBRT: early results from a prospective study in India. *Radiother Oncol* 2019;**133**:S759–60. [https://dx.doi.org/10.1016/s0167-8140\(19\)31814-6](https://dx.doi.org/10.1016/s0167-8140(19)31814-6)
400. Teramoto K, Kawamura T, Takamatsu S, Nakamura N, Kudo A, Noguchi N, *et al.* Laparoscopic and thoracoscopic approaches for the treatment of hepatocellular carcinoma. *Am J Surg* 2005;**189**:474–8.
401. Toro A, Pulvirenti E, Palermo F, Di Carlo I. Health-related quality of life in patients with hepatocellular carcinoma after hepatic resection, transcatheter arterial chemoembolization, radiofrequency ablation or no treatment. *Surg Oncol* 2012;**21**:e23–30. <https://dx.doi.org/10.1016/j.suronc.2011.10.005>
402. Trotschel H. Hepatocellular carcinoma: early MR imaging and monitoring of progression in irreversible electroporation. *Rofo* 2016;**188**:910–1. <http://dx.doi.org/10.1055/s-0036-1585605>
403. Ueno M, Takabatake H, Kayahara T, Morimoto Y, Mizuno M. Stereotactic body radiotherapy versus radiofrequency ablation for single small hepatocellular carcinoma: a propensity score matching analysis. *Hepatology* 2020;**72**:708A–9A.

404. Utsunomiya T, Shimada M, Kudo M, Ichida T, Matsui O, Izumi N, *et al.* Nationwide study of 4741 patients with non-B non-C hepatocellular carcinoma with special reference to the therapeutic impact. *Ann Surg* 2014;**259**:336–45. <http://dx.doi.org/10.1097/SLA.0b013e31829291e9>
405. Vietti Violi N, Duran R, Guiu B, Aube C, Bize PE, Pache I, *et al.* Prospective randomized controlled trial comparing efficacy of microwave ablation and radiofrequency ablation for the treatment of hepatocellular carcinoma in patients with a chronic liver disease. *Cardiovasc Intervent Radiol* 2017;**40**:S133. <http://dx.doi.org/10.1007/s00270-017-1725-y>
406. Vietti VN, Duran R, Guiu B, Aube C, Cercueil JP, Bize P, *et al.* Microwave ablation and radiofrequency ablation for the treatment of hepatocellular carcinoma: result of the first prospective randomized controlled trial. *Ann Oncol* 2017;**28**:iii6. <http://dx.doi.org/10.1093/annonc/mdx263.013>
407. Vitale A, D'Amico F, Gringeri E, Zanus G, Cillo U. Laparoscopic ablation and salvage transplantation in patients with hepatocellular carcinoma. *Transplantation* 2012;**94**:637.
408. Vitali GC, Laurent A, Terraz S, Majno P, Buchs NC, Rubbia-Brandt L, *et al.* Minimally invasive surgery versus percutaneous radio frequency ablation for the treatment of single small ( $\leq 3$  cm) hepatocellular carcinoma: a case-control study. *Surg Endosc* 2016;**30**:2301–7. <https://dx.doi.org/10.1007/s00464-015-4295-6>
409. Wang S, Huang J, Zhang J, Hu X, Zhu W, Hu L, *et al.* Combination treatment of hepatic carcinoma with transcatheter hepatic arterial chemoembolization and percutaneous hepatic cryoablation. *Technol Cancer Res Treat* 2007;**6**:524–7.
410. Wang JH, Wang CC, Hung CH, Chen CL, Lu SN. Survival comparison between surgical resection and radiofrequency ablation for patients in BCLC very early/early stage hepatocellular carcinoma. *J Hepatol* 2012;**56**:412–8. <https://dx.doi.org/10.1016/j.jhep.2011.05.020>
411. Wang WH, Wang Z, Wu JX, Zhang T, Rong WQ, Wang LM, *et al.* Survival benefit with IMRT following narrow-margin hepatectomy in patients with hepatocellular carcinoma close to major vessels. *Liver Int* 2015;**35**:2603–10. <https://dx.doi.org/10.1111/liv.12857>
412. Wang J, Sun W, Liu Q. Comparative study on effect of radiofrequency ablation and operation for the treatment of small hepatocellular carcinoma. *Cancer Res Clin* 2015;**27**:620–3. <http://dx.doi.org/10.3760/cma.j.issn.1006-9801.2015.09.012>
413. Wang G, Zhang W, Tan Y, Jiang L, Yang J, Yang J, *et al.* The risk factors for long-term survival outcome in solitary hepatocellular carcinoma up to 2cm: propensity score matching analysis in a population cohort with a high rate of HBV infection. *Int J Surg* 2019;**72**:1–6. <https://dx.doi.org/10.1016/j.ijssu.2019.10.006>
414. Wang L, Deng D, Deng C. Clinical efficacy and long-term prognosis of laparoscopic liver resection and radiofrequency ablation for small hepatocellular carcinoma. *Int J Clin Exp Med* 2020;**13**:7349–56.
415. Wei J, Li J, Fan W. Unresectable hepatocellular carcinoma: transcatheter arterial chemoembolization combined with microwave ablation versus combined with cryoablation. *Gut* 2020;**69**:A76–7. <https://dx.doi.org/10.1136/gutjnl-2020-IDDF.146>
416. Wigg AJ, Muller KR, Teo M, Rodgers A, Mounkley D, Chinnaratha MA, *et al.* Stereotactic ablative body radiotherapy for early hepatocellular carcinoma: a promising early South Australian experience. *J Gastroenterol Hepatol* 2017;**32**:111.
417. Wiggermann P, Puls R, Vasilij A, Sieron D, Schreyer AG, Jung EM, *et al.* Thermal ablation of unresectable liver tumors: factors associated with partial ablation and the impact on long-term survival. *Med Sci Monit* 2012;**18**:CR88–92.

418. Wong TC, Lee VH, Law AL, Pang HH, Lam KO, Lau V, *et al.* Prospective study of stereotactic body radiation therapy for hepatocellular carcinoma on waitlist for liver transplant. *Hepatology* 2021;**74**:2580–94. <https://dx.doi.org/10.1002/hep.31992>
419. Wu D, Yang Y, Chen J, Cai H, Duan Y, Sun D. Three different ways of treating primary hepatocellular carcinoma at an early stage: a prospective comparative study. *Gastroenterol Res Pract* 2020;**2020**:7802498. <https://dx.doi.org/10.1155/2020/7802498>
420. Xie L, Cao F, Qi H, Song Z, Shen L, Chen S, *et al.* Efficacy and safety of CT-guided percutaneous thermal ablation for hepatocellular carcinoma adjacent to the second porta hepatis. *Int J Hyperthermia* 2019;**36**:1122–8. <https://dx.doi.org/10.1080/02656736.2019.1684575>
421. Xu KC, Niu LZ, Zhou Q, Hu YZ, Guo DH, Liu ZP, *et al.* Sequential use of transarterial chemoembolization and percutaneous cryosurgery for hepatocellular carcinoma. *World J Gastroenterol* 2009;**15**:3664–9.
422. Xu W, Kim BW, Park YK, Cheong JY, Cho SW, Wang HJ. Comparison of hepatic resection and radiofrequency ablation for solitary hepatocellular carcinoma smaller than 4 cm. *HPB (Oxford)* 2014;**16**:459.
423. Xu XS, Chen W, Miao RC, Zhou YY, Wang ZX, Zhang LQ, *et al.* Survival analysis of hepatocellular carcinoma: a comparison between young patients and aged patients. *Chin Med J (Engl)* 2015;**128**:1793–800. <https://dx.doi.org/10.4103/0366-6999.159356>
424. Yamao T, Imai K, Yamashita YI, Kaida T, Nakagawa S, Mima K, *et al.* Surgical treatment strategy for hepatocellular carcinoma in patients with impaired liver function: hepatic resection or radiofrequency ablation? *HPB (Oxford)* 2018;**20**:244–50. <https://dx.doi.org/10.1016/j.hpb.2017.08.031>
425. Yamashita S, Katagiri S, Ariizumi SI, Ohmori A, Takahashi Y, Namoto S, *et al.* Hepatectomy vs. radio frequency ablation as first-line treatment for a small hepatocellular carcinoma. *J Hepatobiliary Pancreat Sci* 2017;**24**:A242.
426. Yamashita S, Ariizumi S, Otera YK, Omori A, Kato T, Nemoto S, *et al.* Clinical outcome by tumor diameter in initial treatment for small hepatocellular carcinoma. *HPB (Oxford)* 2019;**21**:S371–S2. <http://dx.doi.org/10.1016/j.hpb.2019.10.2009>
427. Yamazaki O, Oka H, Manabe T, Kioka K, Kurai O, Murata K, *et al.* Comparison of the long-term results between surgical resection and local ablation for Child–Pugh A patients with a single hepatocellular carcinoma of 5 cm or less in size: a single center study. *Kanzo* 2009;**50**:173–84. <http://dx.doi.org/10.2957/kanzo.50.173>
428. Yang ZX, Wang D, Wang G, Zhang QH, Liu JM, Peng P, *et al.* Clinical study of recombinant adenovirus-p53 combined with fractionated stereotactic radiotherapy for hepatocellular carcinoma. *J Cancer Res Clin Oncol* 2010;**136**:625–30. <https://dx.doi.org/10.1007/s00432-009-0701-6>
429. Ye X, Ge Z, Fei X, Wang S, Cheng Y, Chen X, *et al.* Clinical study on HIFU combined with Jinlong capsules in treating 54 cases of primary liver cancer. *Chin J Clin Oncol* 2008;**35**:372–4.
430. Yi Y, Zhang Y, Wei Q, Zhao L, Han J, Song Y, *et al.* Radiofrequency ablation or microwave ablation combined with transcatheter arterial chemoembolization in treatment of hepatocellular carcinoma by comparing with radiofrequency ablation alone. *Chin J Cancer Res* 2014;**26**:112–8. <https://doi.org/10.3978/j.issn.1000-9604.2014.02.09>
431. Yohji H, Aikata H, Kimura T, Kakizawa H, Kenjo M, Awai K, *et al.* Pilot study of stereotactic body radiation therapy combined with transcatheter arterial chemoembolization for small hepatocellular carcinoma. *Hepatology* 2012;**56**:621A.

432. Yoon SM, Ryoo BY, Lee SJ, Kim JH, Shin JH, An JH, *et al.* Efficacy and safety of transarterial chemoembolization plus external beam radiotherapy vs sorafenib in hepatocellular carcinoma with macroscopic vascular invasion a randomized clinical trial. *JAMA Oncol* 2018;**4**:661–9. <https://dx.doi.org/10.1001/jamaoncol.2017.5847>
433. Yu WB, Wang WH, Rong WQ, Wang LM, Xu Q, Wu F, *et al.* Adjuvant radiotherapy in centrally located hepatocellular carcinomas after hepatectomy with narrow margin (<1 cm): a prospective randomized study. *J Am Coll Surg* 2014;**218**:381–92. <https://dx.doi.org/10.1016/j.jamcollsurg.2013.11.030>
434. Yun WK, Choi MS, Yoo BC, Lee JH, Koh KC, Paik SW, *et al.* Long-term outcome after surgery is superior to that after radiofrequency ablation in Child–Pugh class A patients with single small hepatocellular carcinoma. *Hepatology* 2009;**50**:1081A. <http://dx.doi.org/10.1002/hep.23307>
435. Zhang J, Zhang C, Chen S, Lai J, Zheng W, Wang X. Effect of stereotactic radiotherapy (SRT) combined with percutaneous ethanol injection therapy (PEIT) for small hepatocarcinoma. *Chin J Clin Oncol* 2008;**35**:852–4.
436. Zheng L, Zhang CH, Lin JY, Song CL, Qi XL, Luo M. Comparative effectiveness of radiofrequency ablation vs. surgical resection for patients with solitary hepatocellular carcinoma smaller than 5 cm. *Front Oncol* 2020;**10**:399. <https://dx.doi.org/10.3389/fonc.2020.00399>



# Appendix 1 Search strategies

The MEDLINE search strategies can be found in [Appendix 1.1–1.4](#), along with a list of further databases and resources searched. All other search strategies can be found in [Report Supplementary Material 1](#).

The terms used in all search strategies build upon those used in the searches to inform a previous systematic review on SIRT therapies for hepatocellular carcinoma:

Walton M, Wade R, Claxton L, Sharif-Hurst S, Harden M, Patel J, *et al.* Selective internal radiation therapies for unresectable early-, intermediate- or advanced-stage hepatocellular carcinoma: systematic review, network meta-analysis and economic evaluation. *Health Technol Assess* 2020;24(48)

## Appendix 1.1 Search strategies for identification of randomised controlled trials

The following databases were searched:

- MEDLINE ALL (Ovid)
- Embase (Ovid)
- Cochrane Central Register of Controlled Trials (CENTRAL) (Wiley)
- Science Citation Index (Web of Science)
- Cochrane Database of Systematic Reviews (CDSR) (Wiley)
- Database of Abstracts of Reviews of Effects (DARE) (CRD databases)
- International Health Technology Assessment database
- Epistemonikos
- International Prospective Register of Systematic Reviews (PROSPERO)
- ClinicalTrials.gov
- European Union Clinical Trials Register.

The MEDLINE search strategy can be found below. See [Report Supplementary Material 1](#) for all other search strategies.

### MEDLINE ALL

(includes: epub ahead of print, in-process and other non-indexed citations, Ovid MEDLINE Daily and Ovid MEDLINE)

via Ovid <http://ovidsp.ovid.com/>

Date range: 1946 to 1 February 2021

Date searched: 3 February 2021

Records retrieved: 2303

The MEDLINE strategy below includes a search filter to limit retrieval to RCTs using the Cochrane Highly Sensitive Search Strategy for identifying randomised trials in MEDLINE: sensitivity-maximising version (2008 revision); Ovid format.

Lefebvre C, Glanville J, Briscoe S, Littlewood A, Marshall C, Metzendorf M-I, Noel-Storr A, Rader T, Shokraneh F, Thomas J, Wieland LS. Technical Supplement to Chapter 4: Searching for and selecting studies. In: Higgins JPT, Thomas J, Chandler J, Cumpston MS, Li T, Page MJ, Welch VA (eds). *Cochrane*

*Handbook for Systematic Reviews of Interventions* Version 6.2 (updated February 2021). Cochrane, 2021. Available from: [www.training.cochrane.org/handbook](http://www.training.cochrane.org/handbook).

- 1 Carcinoma, Hepatocellular/ (86,979)
- 2 Liver Neoplasms/ (151,355)
- 3 ((liver or hepatocellular or hepato-cellular or hepatic\$) adj3 (carcinoma\$ or cancer\$ or neoplas\$ or tumour\$ or tumor\$ or malign\$)).ti,ab. (150,762)
- 4 (hepatocellularcarcinoma\$ or hepatocarcinoma\$ or hepato-carcinoma\$).ti,ab. (4183)
- 5 hepatoma\$.ti,ab. (28,611)
- 6 HCC.ti,ab. (58,929)
- 7 or/1-6 (234,592)
- 8 Neoplasm Staging/ (177,611)
- 9 (small\$ or early or earlystage?).ti,ab. (3,163,695)
- 10 (((BCLC or Barcelona-Clinic Liver Cancer) adj3 ("0" or A or A1 or A2 or A3 or A4)) or BCLC0-A).ti,ab. (578)
- 11 ("1" or "2" or "3" or one or two or three) adj (cm\$ or centimet\$).ti,ab. (77,157)
- 12 (1cm\$ or 2cm\$ or 3cm\$).ti,ab. (4783)
- 13 ((carcinoma\$ or tumor\$ or tumour\$ or lesion\$ or nodule\$) adj6 (size\$ or diameter\$)).ti,ab. (124,495)
- 14 (eHCC or sHCC).ti,ab. (251)
- 15 or/8-13 (3,424,647)
- 16 14 or (7 and 15) (50,618)
- 17 Radiofrequency Ablation/ (1071)
- 18 Catheter Ablation/ (33,067)
- 19 Radiofrequency Therapy/ (1098)
- 20 ((radiofrequenc\$ or radio frequenc\$) adj3 ablat\$).ti,ab. (20,597)
- 21 RFA.ti,ab. (6924)
- 22 RF ablation.ti,ab. (2496)
- 23 RTA.ti,ab. (2494)
- 24 RFTA.ti,ab. (62)

- 25 or/17-24 (45,285)
- 26 16 and 25 (3125)
- 27 Microwaves/ (17,445)
- 28 (microwave\$ or micro wave\$).ti,ab. (38,436)
- 29 (MWA or MCT or PMCT or PMWA).ti,ab. (7692)
- 30 or/27-29 (47,890)
- 31 16 and 30 (715)
- 32 Laser Therapy/ (38,202)
- 33 (laser\$ adj2 ablat\$).ti,ab. (10,008)
- 34 LTA.ti,ab. (3499)
- 35 or/32-34 (48,883)
- 36 16 and 35 (142)
- 37 High-Intensity Focused Ultrasound Ablation/ (1728)
- 38 Ultrasonic Therapy/ (9714)
- 39 High intensity focus?ed ultrasound.ti,ab. (3137)
- 40 HIFU.ti,ab. (2449)
- 41 or/37-40 (12,933)
- 42 16 and 41 (132)
- 43 Cryosurgery/ (13,169)
- 44 Cryotherapy/ (5214)
- 45 (cryoablat\$ or cryo-ablat\$ or cryotherap\$ or cryo-therap\$ or cryosurg\$ or cryo-surg\$).ti,ab. (14,455)
- 46 or/43-45 (23,464)
- 47 16 and 46 (317)
- 48 Ethanol/ (88,324)
- 49 ((alcohol or ethanol) adj2 (inject\$ or ablat\$)).ti,ab. (5254)
- 50 (PEI or PEIT).ti,ab. (8408)
- 51 or/48-50 (98,037)

## APPENDIX 1

52 16 and 51 (975)

53 Acetic Acid/ (10,269)

54 Acetates/ (39,925)

55 (acetic acid adj2 (inject\$ or ablat\$)).ti,ab. (408)

56 PAI.ti,ab. (14,810)

57 PAAI.ti,ab. (19)

58 or/53-57 (62,615)

59 16 and 58 (132)

60 Electroporation/ (8033)

61 electroporation.ti,ab. (10,705)

62 IRE.ti,ab. (2151)

63 or/60-62 (15,275)

64 16 and 63 (122)

65 ((stereotactic or stereotaxic) adj3 ablat\$).ti,ab. (1271)

66 ((stereotactic or stereotaxic) adj3 (radiotherap\$ or radiation)).ti,ab. (9612)

67 (SABR or SABRT).ti,ab. (803)

68 SBRT.ti,ab. (4238)

69 SABER.ti,ab. (356)

70 or/65-69 (10,729)

71 16 and 70 (428)

72 Ablation Techniques/ (2918)

73 (ablat\$ adj2 (therap\$ or intervention\$ or treatment\$ or technique\$ or method\$ or procedure\$)).ti,ab. (16,748)

74 (ablat\$ adj2 (chemical\$ or thermal\$)).ti,ab. (4065)

75 (ablat\$ adj2 (tumour\$ or tumor\$)).ti,ab. (4083)

76 or/72-75 (24,549)

77 16 and 76 (1788)

- 78 Chemoembolization, Therapeutic/ (5983)
- 79 (chemo-emboli\$ or chemoemboli\$).ti,ab. (8271)
- 80 TACE.ti,ab. (5534)
- 81 cTACE.ti,ab. (144)
- 82 (DEBTACE or DEB-TACE).ti,ab. (243)
- 83 (eluting adj2 bead\$).ti,ab. (624)
- 84 DC bead\$.ti,ab. (108)
- 85 or/78-84 (11,130)
- 86 16 and 85 (3538)
- 87 Embolization, Therapeutic/ (32,829)
- 88 (embolization\$ or embolisation\$ or embolize\$ or embolise\$ or embolizing\$ or embolising\$ or embolotherap\$).ti,ab. (52,479)
- 89 TAE.ti,ab. (2435)
- 90 or/87-89 (63,331)
- 91 16 and 90 (2015)
- 92 ((locoregional or loco-regional) adj2 (therap\$ or intervention\$ or treatment\$ or technique\$ or method\$ or procedure\$)).ti,ab. (3346)
- 93 16 and 92 (531)
- 94 (Therasphere\$ or Thera-sphere\$).ti,ab. (79)
- 95 (SIR-Sphere\$ or SIRSphere\$).ti,ab. (119)
- 96 (QuiremSphere\$ or Quirem-Sphere\$).ti,ab. (4)
- 97 or/94-96 (167)
- 98 16 and 97 (44)
- 99 Microspheres/ (28,670)
- 100 (microsphere\$ or sphere\$).ti,ab. (76,678)
- 101 (microbead\$ or bead\$).ti,ab. (56,354)
- 102 or/99-101 (139,820)
- 103 Yttrium Radioisotopes/ (3105)

- 104 Yttrium/ (3157)
- 105 Yttrium Isotopes/ (709)
- 106 (Yttrium\$ or 90Yttrium\$ or Y90 or Y-90 or 90Y or 90-Y).ti,ab. (9775)
- 107 Holmium/ (904)
- 108 (Holmium\$ or 166Holmium\$ or Ho-166 or Ho166 or 166Ho or 166-Ho).ti,ab. (3496)
- 109 Radiopharmaceuticals/ (51,067)
- 110 or/103-109 (66,136)
- 111 102 and 110 (1871)
- 112 ((radioactiv\$ or radio-activ\$ or radionuclide\$ or radio-nuclide\$ or radioisotope\$ or radio-isotope\$ or radiolabel\$ or radio-label\$ or radiopharmaceutic\$ or radio-pharmaceutic\$) adj2 (sphere\$ or microsphere\$ or bead\$ or microbead\$)).ti,ab. (4168)
- 113 (radiomicrosphere\$ or radio-microsphere\$).ti,ab. (33)
- 114 or/111-113 (5932)
- 115 16 and 114 (315)
- 116 Brachytherapy/ (19,954)
- 117 (brachytherap\$ or brachy-therap\$ or microbrachytherap\$).ti,ab. (18,064)
- 118 or/116-117 (25,595)
- 119 118 and (110 or 112 or 113) (1048)
- 120 16 and 119 (85)
- 121 (radioemboli\$ or radio-emboli\$ or radioembolotherap\$ or radio-embolotherap\$).ti,ab. (1791)
- 122 TARE.ti,ab. (276)
- 123 (internal\$ adj3 (radiation\$ or radiotherap\$ or radio therap\$ or radionuclide\$ or radio-nuclide\$ or radioisotope\$ or radio-isotope\$)).ti,ab. (2446)
- 124 ((intra-arterial\$ or intraarterial\$) adj3 (radiation\$ or radiotherap\$ or radio therap\$ or radionuclide\$ or radio-nuclide\$ or radioisotope\$ or radio-isotope\$)).ti,ab. (284)
- 125 ((intra-arterial\$ or intraarterial\$) adj2 (brachytherap\$ or brachy-therap\$)).ti,ab. (20)
- 126 SIRT.ti,ab. (1519)
- 127 (SIR adj2 (therap\$ or treatment\$)).ti,ab. (88)
- 128 (radiation adj2 (segmentectom\$ or lobectom\$)).ti,ab. (53)

129 or/121-128 (5699)

130 16 and 129 (617)

131 26 or 31 or 36 or 42 or 47 or 52 or 59 or 64 or 71 or 77 (5352)

132 86 or 91 or 93 or 98 or 115 or 120 or 130 (5241)

133 randomized controlled trial.pt. (521,951)

134 controlled clinical trial.pt. (94,049)

135 randomized.ab. (510,387)

136 placebo.ab. (215,580)

137 drug therapy.fs. (2,274,478)

138 randomly.ab. (351,559)

139 trial.ab. (541,682)

140 groups.ab. (2,157,357)

141 133 or 134 or 135 or 136 or 137 or 138 or 139 or 140 (4,916,502)

142 exp animals/ not humans.sh. (4,782,806)

143 141 not 142 (4,274,490)

144 131 and 143 (1485)

145 132 and 143 (1633)

146 144 or 145 (2615)

147 limit 146 to yr="2000 -Current" (2303)

### Key:

/ = subject heading (MeSH heading)

sh = subject heading (MeSH heading)

exp = exploded subject heading (MeSH heading)

\$ = truncation

? = optional wild card character – stands for zero or one characters

ti,ab = terms in title or abstract fields

adj3 = terms within three words of each other (any order)

pt = publication type

fs = floating subheading

**Search strategies for identification of randomised controlled trials of wider radiotherapy techniques (March 2021)**

**MEDLINE ALL**

(includes: epub ahead of print, in-process and other non-indexed citations, Ovid MEDLINE Daily and Ovid MEDLINE)

via Ovid <http://ovidsp.ovid.com/>

Date range: 1946 to 16 March 2021

Date searched: 17 March 2021

Records retrieved: 399

The MEDLINE strategy below includes a search filter to limit retrieval to RCTs using the Cochrane Highly Sensitive Search Strategy for identifying randomised trials in MEDLINE: sensitivity-maximising version (2008 revision); Ovid format.

Lefebvre C, Glanville J, Briscoe S, Littlewood A, Marshall C, Metzendorf M-I, Noel-Storr A, Rader T, Shokraneh F, Thomas J, Wieland LS. Technical Supplement to Chapter 4: Searching for and selecting studies. In: Higgins JPT, Thomas J, Chandler J, Cumpston MS, Li T, Page MJ, Welch VA (eds). *Cochrane Handbook for Systematic Reviews of Interventions* Version 6.2 (updated February 2021). Cochrane, 2021. Available from: [www.training.cochrane.org/handbook](http://www.training.cochrane.org/handbook).

1 Carcinoma, Hepatocellular/ (87,724)

2 Liver Neoplasms/ (152,378)

3 ((liver or hepatocellular or hepato-cellular or hepatic\$) adj3 (carcinoma\$ or cancer\$ or neoplas\$ or tumour\$ or tumor\$ or malign\$)).ti,ab. (151,424)

4 (hepatocellularcarcinoma\$ or hepatocarcinoma\$ or hepato-carcinoma\$).ti,ab. (4202)

5 hepatoma\$.ti,ab. (28,603)

6 HCC.ti,ab. (59,303)

7 or/1-6 (235,478)

8 Neoplasm Staging/ (178,539)

9 (small\$ or early or earlystage?).ti,ab. (3,174,425)

10 (((BCLC or Barcelona-Clinic Liver Cancer) adj3 ("0" or A or A1 or A2 or A3 or A4)) or BCLC0-A).ti,ab. (588)

11 ("1" or "2" or "3" or one or two or three) adj (cm\$ or centimet\$).ti,ab. (77,490)

- 12 (1cm\$ or 2cm\$ or 3cm\$).ti,ab. (4785)
- 13 ((carcinoma\$ or tumor\$ or tumour\$ or lesion\$ or nodule\$) adj6 (size\$ or diameter\$)).ti,ab. (125,055)
- 14 (eHCC or sHCC).ti,ab. (254)
- 15 or/8-13 (3,436,575)
- 16 14 or (7 and 15) (50,862)
- 17 Electrochemotherapy/ (673)
- 18 (electrochemotherap\* or electro-chemotherap\* or electro chemotherap\* or electroporabili?ation).ti,ab. (1115)
- 19 (electric\* adj2 stimulat\* adj2 (therap\* or chemotherap\* or chemo-therap\* or chemo therap\* or treat\*)).ti,ab. (1260)
- 20 or/17-19 (2670)
- 21 histotripsy.ti,ab. (209)
- 22 Radiotherapy/ or Radiotherapy, Conformal/ or Radiotherapy, Intensity-Modulated/ or Radiotherapy, High-Energy/ or Radiotherapy, Image-Guided/ (72,234)
- 23 (radiotherap\* or radiation-therap\* or radiation therap\*).ti,ab. (237,815)
- 24 ((intensity-modulat\* or intensity modulat\* or volumetric-modulat\* or volumetric modulat\*) adj4 (arc therap\* or arc-therap\*)).ti,ab. (2469)
- 25 (helical\* adj4 tomotherap\*).ti,ab. (1214)
- 26 or/22-25 (267,755)
- 27 Proton Therapy/ (3960)
- 28 (proton\* adj4 therap\*).ti,ab. (7126)
- 29 or/27-28 (8349)
- 30 20 or 21 or 26 or 29 (275,176)
- 31 16 and 30 (1887)
- 32 randomized controlled trial.pt. (525,223)
- 33 controlled clinical trial.pt. (94,097)
- 34 randomized.ab. (512,974)
- 35 placebo.ab. (216,151)
- 36 drug therapy.fs. (2,290,533)

37 randomly.ab. (353,254)

38 trial.ab. (543,763)

39 groups.ab. (2,167,571)

40 or/32-39 (4,942,795)

41 31 and 40 (496)

42 exp animals/ not humans.sh. (4,800,681)

43 41 not 42 (478)

44 limit 43 to yr="2000 -Current" (399)

**Key:**

/ = subject heading (MeSH heading)

sh = subject heading (MeSH heading)

exp = exploded subject heading (MeSH heading)

\$ = truncation

? = optional wild card character – stands for zero or one characters

ti,ab = terms in title or abstract fields

adj3 = terms within three words of each other (any order)

pt = publication type

fs = floating subheading

**Appendix 1.2 Search strategies for identification of non-randomised studies where randomised controlled trial evidence was not available**

The following databases were searched:

- MEDLINE ALL (Ovid)
- Embase (Ovid)
- Cochrane Central Register of Controlled Trials (CENTRAL) (Wiley)
- Science Citation Index (Web of Science).

The MEDLINE search strategy can be found below. See [Report Supplementary Material 1](#) for all other search strategies.

**MEDLINE ALLw**

(includes: epub ahead of print, in-process and other non-indexed citations, Ovid MEDLINE Daily and Ovid MEDLINE)

via Ovid <http://ovidsp.ovid.com/>

Date range: 1946 to 27 July 2021

Date searched: 28 July 2021

Records retrieved: 1139

1 Carcinoma, Hepatocellular/ (90,761)

2 Liver Neoplasms/ (156,595)

3 ((liver or hepatocellular or hepato-cellular or hepatic\$) adj3 (carcinoma\$ or cancer\$ or neoplas\$ or tumour\$ or tumor\$ or malign\$)).ti,ab. (156,050)

4 (hepatocellularcarcinoma\$ or hepatocarcinoma\$ or hepato-carcinoma\$).ti,ab. (4256)

5 hepatoma\$.ti,ab. (28,811)

6 HCC.ti,ab. (61,752)

7 or/1-6 (241,089)

8 Neoplasm Staging/ (182,051)

9 (small\$ or early or earlystage?).ti,ab. (3,247,367)

10 (((BCLC or Barcelona-Clinic Liver Cancer) adj3 ("0" or A or A1 or A2 or A3 or A4)) or BCLCO-A).ti,ab. (617)

11 ("1" or "2" or "3" or one or two or three) adj (cm\$ or centimet\$).ti,ab. (79,017)

12 (1cm\$ or 2cm\$ or 3cm\$).ti,ab. (4905)

13 ((carcinoma\$ or tumor\$ or tumour\$ or lesion\$ or nodule\$) adj6 (size\$ or diameter\$)).ti,ab. (128,094)

14 (eHCC or sHCC).ti,ab. (265)

15 or/8-13 (3,515,054)

16 14 or (7 and 15) (52,166)

17 High-Intensity Focused Ultrasound Ablation/ (1845)

18 Ultrasonic Therapy/ (9862)

19 High intensity focus?ed ultrasound.ti,ab. (3223)

20 HIFU.ti,ab. (2524)

21 or/17-20 (13,202)

22 16 and 21 (133)

## APPENDIX 1

23 Cryosurgery/ (13,402)

24 Cryotherapy/ (5359)

25 (cryoablat\$ or cryo-ablat\$ or cryotherap\$ or cryo-therap\$ or cryosurg\$ or cryo-surg\$).ti,ab. (14,771)

26 or/23-25 (23,935)

27 16 and 26 (323)

28 Electroporation/ (8248)

29 electroporation.ti,ab. (10,923)

30 IRE.ti,ab. (2244)

31 or/28-30 (15,583)

32 16 and 31 (133)

33 ((stereotactic or stereotaxic) adj3 ablat\$).ti,ab. (1364)

34 ((stereotactic or stereotaxic) adj3 (radiotherap\$ or radiation)).ti,ab. (10,149)

35 (SABR or SABRT).ti,ab. (859)

36 SBRT.ti,ab. (4543)

37 SABER.ti,ab. (421)

38 or/33-37 (11,398)

39 16 and 38 (456)

40 Electrochemotherapy/ (698)

41 (electrochemotherap\* or electro-chemotherap\* or electro chemotherap\* or electropermeabili?ation).ti,ab. (1153)

42 (electric\* adj2 stimulat\* adj2 (therap\* or chemotherap\* or chemo-therap\* or chemo therap\* or treat\*)).ti,ab. (1304)

43 40 or 41 or 42 (2757)

44 16 and 43 (25)

45 histotripsy.ti,ab. (230)

46 16 and 45 (7)

47 Radiotherapy, Conformal/ or Radiotherapy, Intensity-Modulated/ or Radiotherapy, High-Energy/ or Radiotherapy, Image-Guided/ (30,927)

48 ((radiotherap\* or radiation-therap\* or radiation therap\*) adj3 (conformal or intensity-modulat\* or intensity modulat\* or high-energy or high energy)).ti,ab. (14,403)

49 ((intensity-modulat\* or intensity modulat\* or volumetric-modulat\* or volumetric modulat\*) adj4 (arc therap\* or arc-therap\*)).ti,ab. (2595)

50 (helical\* adj4 tomotherap\*).ti,ab. (1237)

51 47 or 48 or 49 or 50 (36,559)

52 16 and 51 (274)

53 Proton Therapy/ (4266)

54 (proton\* adj4 therap\*).ti,ab. (7420)

55 53 or 54 (8702)

56 16 and 55 (106)

57 22 or 27 or 32 or 39 or 44 or 46 or 52 or 56 (1292)

58 exp animals/ not humans.sh. (4,866,074)

59 57 not 58 (1226)

60 limit 59 to yr="2000-Current" (1139)

### Key:

/ = subject heading (MeSH heading)

sh = subject heading (MeSH heading)

exp = exploded subject heading (MeSH heading)

\$ = truncation

\* = truncation

? = optional wild card character – stands for zero or one characters

ti,ab = terms in title or abstract fields

adj3 = terms within three words of each other (any order)

### Appendix 1.3 Search strategies for identification of non-randomised studies where additional evidence could plausibly change the network meta-analysis result, as identified by the threshold analysis

The following databases were searched:

- MEDLINE ALL (Ovid)
- Embase (Ovid)
- Cochrane Central Register of Controlled Trials (CENTRAL) (Wiley)
- Science Citation Index (Web of Science)

The MEDLINE search strategy can be found below. See [Report Supplementary Material 1](#) for all other search strategies.

#### MEDLINE ALL

(includes: epub ahead of print, in-process and other non-indexed citations, Ovid MEDLINE Daily and Ovid MEDLINE)

via Ovid <http://ovidsp.ovid.com/>

Date range: 1946 to 23 August 2021

Date searched: 24 August 2021

Records retrieved: 2539

1 Carcinoma, Hepatocellular/ (91,350)

2 Liver Neoplasms/ (157,448)

3 ((liver or hepatocellular or hepato-cellular or hepatic\$) adj3 (carcinoma\$ or cancer\$ or neoplas\$ or tumour\$ or tumor\$ or malign\$)).ti,ab. (156,958)

4 (hepatocellularcarcinoma\$ or hepatocarcinoma\$ or hepato-carcinoma\$).ti,ab. (4266)

5 hepatoma\$.ti,ab. (28,844)

6 HCC.ti,ab. (62,236)

7 or/1-6 (242,196)

8 Neoplasm Staging/ (182,692)

9 (small\$ or early or earlystage?).ti,ab. (3,261,749)

10 (((BCLC or Barcelona-Clinic Liver Cancer) adj3 ("0" or A or A1 or A2 or A3 or A4)) or BCLC0-A).ti,ab. (628)

11 ("1" or "2" or "3" or one or two or three) adj (cm\$ or centimet\$).ti,ab. (79,327)

12 (1cm\$ or 2cm\$ or 3cm\$).ti,ab. (4920)

13 ((carcinoma\$ or tumor\$ or tumour\$ or lesion\$ or nodule\$) adj6 (size\$ or diameter\$)).ti,ab. (128,670)

- 14 (eHCC or sHCC).ti,ab. (268)
- 15 or/8-13 (3,530,492)
- 16 14 or (7 and 15) (52,397)
- 17 Laser Therapy/ (38,954)
- 18 (laser\$ adj2 ablat\$).ti,ab. (10,374)
- 19 LTA.ti,ab. (3598)
- 20 or/17-19 (49,955)
- 21 Radiofrequency Ablation/ (1620)
- 22 Catheter Ablation/ (34,863)
- 23 Radiofrequency Therapy/ (1141)
- 24 ((radiofrequenc\$ or radio frequenc\$) adj3 ablat\$).ti,ab. (21,317)
- 25 RFA.ti,ab. (7252)
- 26 RF ablation.ti,ab. (2534)
- 27 RTA.ti,ab. (2555)
- 28 RFTA.ti,ab. (63)
- 29 or/21-28 (47,522)
- 30 Microwaves/ (18,128)
- 31 (microwave\$ or micro wave\$).ti,ab. (39,869)
- 32 (MWA or MCT or PMCT or PMWA).ti,ab. (8015)
- 33 or/30-32 (49,569)
- 34 Carcinoma, Hepatocellular/su (13,662)
- 35 Liver Neoplasms/su (27,831)
- 36 Hepatectomy/ (31,584)
- 37 Surgical Procedures, Operative/ (56,264)
- 38 ((surgical\$ or surger\$ or operat\$ or resect\$) adj6 (carcinoma\$ or tumor\$ or tumour\$ or lesion\$ or nodule\$ or neoplasm\$ or liver\$ or lobe\$)).ti,ab. (257,196)
- 39 (hepatectom\$ or hemi-hepatectom\$ or hemihepatectom\$ or lobectom\$ or microlobectom\$ or microlobectom\$ or segmentectom\$ or trisegmentectom\$).ti,ab. (46,091)

40 ((carcinoma\$ or tumor\$ or tumour\$ or lesion\$ or nodule\$ or neoplasm\$ or liver\$ or lobe\$) adj4 (excis\$ or remov\$ or dissect\$)).ti,ab. (71,760)

41 or/34-40 (396,793)

42 20 and (29 or 33 or 41) (3589)

43 29 and (20 or 33 or 41) (8135)

44 33 and (20 or 29 or 41) (2905)

45 41 and (20 or 29 or 33) (9295)

46 16 and (42 or 43 or 44 or 45) (2696)

47 exp animals/ not humans/ (4,877,412)

48 46 not 47 (2619)

49 limit 48 to yr="2000-Current" (2539)

**Key:**

/ = indexing term (Medical Subject Heading: MeSH)

/su = indexing term with subheading for surgery

exp = exploded indexing term (MeSH)

\$ = truncation

ti,ab = terms in either title or abstract fields

adj3 = terms within three words of each other (any order).

## Appendix 1.4 Search strategies for identification of economic studies

The following databases were searched:

- MEDLINE ALL (Ovid)
- Embase (Ovid)
- EconLit (Ovid)
- NHS Economic Evaluations Database (CRD databases).

The MEDLINE search strategy can be found below. See [Report Supplementary Material 1](#) for all other search strategies.

### **Ovid MEDLINE(R) ALL**

(includes epub ahead of print, in-process and other non-indexed citations, Ovid MEDLINE Daily and Ovid MEDLINE)

via Ovid <http://ovidsp.ovid.com/>

Date range searched: 1946 to 12 May 2021

Date searched: 13 May 2021

Records retrieved: 181

The MEDLINE strategy below (lines 17–24) includes a narrow search filter to limit retrieval to economic studies. The filter was designed by the *Canadian Journal of Health Technologies* (CADTH).

Strings attached: CADTH database search filters [Internet]. Ottawa: CADTH; 2016. [cited 2021 05 13]. Available from: <https://www.cadth.ca/resources/finding-evidence/strings-attached-cadths-database-search-filters#narrow>

- 1 Carcinoma, Hepatocellular/ (88,831)
- 2 Liver Neoplasms/ (153,869)
- 3 ((liver or hepatocellular or hepato-cellular or hepatic\$) adj3 (carcinoma\$ or cancer\$ or neoplas\$ or tumour\$ or tumor\$ or malign\$)).ti,ab. (153,407)
- 4 (hepatocellularcarcinoma\$ or hepatocarcinoma\$ or hepato-carcinoma\$).ti,ab. (4221)
- 5 hepatoma\$.ti,ab. (28,694)
- 6 HCC.ti,ab. (60,357)
- 7 or/1-6 (237,841)
- 8 Neoplasm Staging/ (179,919)
- 9 (small\$ or early or earlystage?).ti,ab. (3,206,080)
- 10 (((BCLC or Barcelona-Clinic Liver Cancer) adj3 ("0" or A or A1 or A2 or A3 or A4)) or BCLCO-A).ti,ab. (604)
- 11 (("1" or "2" or "3" or one or two or three) adj (cm\$ or centimet\$)).ti,ab. (78,139)
- 12 (1cm\$ or 2cm\$ or 3cm\$).ti,ab. (4828)
- 13 ((carcinoma\$ or tumor\$ or tumour\$ or lesion\$ or nodule\$) adj6 (size\$ or diameter\$)).ti,ab. (126,376)
- 14 (eHCC or sHCC).ti,ab. (257)
- 15 or/8-13 (3,470,534)
- 16 14 or (7 and 15) (51,398)
- 17 \*economics/ (10,739)
- 18 exp \*"costs and cost analysis"/ (74,182)
- 19 (economic adj2 model\*).mp. (13,712)

20 (cost minimi\* or cost-utilit\* or health utilit\* or economic evaluation\* or economic review\* or cost outcome or cost analys?s or economic analys?s or budget\* impact analys?s).ti,ab,kf,kw. (35,068)

21 (cost-effective\* or pharmacoeconomic\* or pharmaco-economic\* or cost-benefit or costs).ti,kf,kw. (76,616)

22 (life year or life years or qaly\* or cost-benefit analys?s or cost-effectiveness analys?s).ab,kf,kw. (32,546)

23 (cost or economic\*).ti,kf,kw. and (costs or cost-effectiveness or markov).ab. (61,298)

24 or/17-23 (187,991)

25 16 and 24 (199)

26 exp animals/ not humans.sh. (4,823,832)

27 25 not 26 (199)

28 limit 27 to yr="2000-Current" (181)

**Key:**

/ or.sh. = indexing term (Medical Subject Heading: MeSH)

exp = exploded indexing term (MeSH)

\$ or \* = truncation

ti,ab = terms in either title or abstract fields

kw,kf = terms in keyword or keyfield field

adj3 = terms within three words of each other (any order).

mp = multipurpose field

? = replaces or adds up to one additional character

## Appendix 2 Studies excluded at full paper stage with rationale (randomised controlled trial searches)

Study	Reason for exclusion
An, 2021 <sup>115</sup>	Not a RCT
Chang, 2018 <sup>116</sup>	No relevant intervention/comparison
Cho, 2014 <sup>117</sup>	Not a RCT
Chong, 2017a <sup>118</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Chong, 2017b <sup>119</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Chong, 2018 <sup>120</sup>	Duplicate report
Chong, 2020 <sup>121</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Crocetti, 2018 <sup>122</sup>	Not a RCT
Di Costanzo, 2011a <sup>123</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Di Costanzo, 2011b <sup>124</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Di Costanzo, 2011c <sup>125</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Di Costanzo, 2013 <sup>126</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Di Costanzo, 2015 <sup>127</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Duan, 2011 <sup>128</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
DuBay, 2011 <sup>129</sup>	Not a RCT
Fan, 2019 <sup>130</sup>	Not a RCT
Fang, 2005 <sup>131</sup>	Not a RCT
Ferrer Puchol, 2011 <sup>132</sup>	Not a RCT
Filippiadis, 2021 <sup>133</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Fong, 2016 <sup>134</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Frangakis, 2010 <sup>135</sup>	Not a RCT
Fukushima, 2015 <sup>136</sup>	No relevant intervention/comparison
Gerunda, 2000 <sup>137</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Giorgio, 2010a <sup>138</sup>	Duplicate report
Giorgio, 2010b <sup>139</sup>	Duplicate report
Guo, 2005 <sup>140</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Guo, 2013 <sup>141</sup>	Not a RCT
Ha, 2016 <sup>142</sup>	Not a RCT
Hayes, 2008 <sup>143</sup>	Not a RCT
He, 2018 <sup>144</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Hong, 2005 <sup>145</sup>	Not a RCT

continued

Study	Reason for exclusion
Hsiao, 2020 <sup>146</sup>	Not a RCT
Hsu, 2012 <sup>147</sup>	Not a RCT
Huang, 2009 <sup>148</sup>	Not early HCC patients ( $\leq$ 3 cm tumour)
Huang, 2020 <sup>149</sup>	Not a RCT
Hung, 2011 <sup>150</sup>	Not a RCT
Huo, 2003 <sup>82</sup>	Not a RCT
Hyun, 2016 <sup>151</sup>	Not a RCT
Iida, 2014 <sup>152</sup>	Not a RCT
Ikeda, 2001 <sup>153</sup>	Not a RCT
Imai, 2013 <sup>154</sup>	Not a RCT
Jiang, 2015 <sup>155</sup>	Not a RCT
Jiang, 2017a <sup>156</sup>	Not a RCT
Jiang, 2017b <sup>157</sup>	Not early HCC patients ( $\leq$ 3 cm tumour)
Kaibori, 2012a <sup>158</sup>	Not early HCC patients ( $\leq$ 3 cm tumour)
Kaibori, 2012b <sup>159</sup>	Not early HCC patients ( $\leq$ 3 cm tumour)
Kayali, 2013 <sup>160</sup>	Not early HCC patients ( $\leq$ 3 cm tumour)
Kim, 2019 <sup>161</sup>	Not a RCT
Kim, 2021 <sup>162</sup>	Duplicate report
Kitamoto, 2003 <sup>163</sup>	Not a RCT
Kiyoshi, 2010 <sup>164</sup>	Duplicate report
Kobayashi, 2007 <sup>165</sup>	No relevant intervention/comparison
Koda, 2000 <sup>166</sup>	Duplicate report
Koh, 2015 <sup>167</sup>	Not early HCC patients ( $\leq$ 3 cm tumour)
Kong, 2014 <sup>168</sup>	Not a RCT
Lai, 2016 <sup>169</sup>	Not a RCT
Lambert, 2020 <sup>170</sup>	Not early HCC patients ( $\leq$ 3 cm tumour)
Lee, 2009 <sup>171</sup>	Not a RCT
Lee, 2018a <sup>172</sup>	Not early HCC patients ( $\leq$ 3 cm tumour)
Lee, 2018b <sup>173</sup>	Not a RCT
Li, 2007 <sup>174</sup>	Not early HCC patients ( $\leq$ 3 cm tumour)
Li, 2011 <sup>175</sup>	Not early HCC patients ( $\leq$ 3 cm tumour)
Li, 2012 <sup>176</sup>	Not early HCC patients ( $\leq$ 3 cm tumour)
Liao, 2017 <sup>177</sup>	No relevant intervention/comparison
Lin, 2005 <sup>178</sup>	Duplicate report
Lin, 2007a <sup>179</sup>	Not a RCT
Lin, 2007b <sup>180</sup>	Not a RCT
Liu, 2012 <sup>181</sup>	Not early HCC patients ( $\leq$ 3 cm tumour)

Study	Reason for exclusion
Liu, 2015 <sup>182</sup>	Not early HCC patients ( $\leq$ 3 cm tumour)
Liu, 2016 <sup>183</sup>	Not early HCC patients ( $\leq$ 3 cm tumour)
Liu, 2019 <sup>184</sup>	No relevant intervention/comparison
Liu, 2020 <sup>185</sup>	Not a RCT
Lu, 2006 <sup>186</sup>	Not early HCC patients ( $\leq$ 3 cm tumour)
Lu, 2008 <sup>187</sup>	Not early HCC patients ( $\leq$ 3 cm tumour)
Luo, 2005 <sup>188</sup>	Not a RCT
Ma, 2019 <sup>189</sup>	Not early HCC patients ( $\leq$ 3 cm tumour)
Maeda, 2003 <sup>190</sup>	Not early HCC patients ( $\leq$ 3 cm tumour)
Masuda, 2007 <sup>191</sup>	Not a RCT
Mbalisike, 2015 <sup>192</sup>	Not early HCC patients ( $\leq$ 3-cm tumour)
Meniconi, 2015 <sup>193</sup>	Not a RCT
Meyer, 2013 <sup>194</sup>	Not early HCC patients ( $\leq$ 3 cm tumour)
Mohamed, 2018 <sup>195</sup>	Not early HCC patients ( $\leq$ 3 cm tumour)
Mornex, 2007a <sup>196</sup>	Not a RCT
Mornex, 2007b <sup>197</sup>	Duplicate report
Murakami, 2007 <sup>198</sup>	Not a RCT
Ng, 2017 <sup>73</sup>	Duplicate report
Ni, 2007 <sup>199</sup>	Not early HCC patients ( $\leq$ 3 cm tumour)
NIHR Horizon Scanning Centre <sup>200</sup>	Not a RCT
Nouso, 2016 <sup>201</sup>	No relevant intervention/comparison
Ohmoto, 2006 <sup>202</sup>	Not a RCT
Ohmoto, 2009 <sup>203</sup>	Not a RCT
Olschewski, 2002 <sup>204</sup>	Duplicate report
Paik, 2016 <sup>205</sup>	Not a RCT
Panaro, 2014 <sup>206</sup>	Not a RCT
Park, 2020 <sup>207</sup>	No relevant intervention/comparison
Peng, 2008 <sup>208</sup>	Not a RCT
Petrowsky, 2008 <sup>209</sup>	Not a RCT
Pompili, 2013 <sup>210</sup>	Not a RCT
Riaz, 2009 <sup>211</sup>	Not a RCT
Roche, 2002 <sup>212</sup>	Not a RCT
Ryu, 2017 <sup>213</sup>	Not a RCT
Santambrogio, 2009 <sup>214</sup>	Not a RCT
Shen, 2018 <sup>215</sup>	Not early HCC patients ( $\leq$ 3 cm tumour)
Sherman, 2006 <sup>216</sup>	Not a RCT

continued

Study	Reason for exclusion
Shi, 2014 <sup>217</sup>	Not a RCT
Shibata, 2006 <sup>218</sup>	No relevant intervention/comparison
Sun, 2020 <sup>219</sup>	Not a RCT
Tashiro, 2011 <sup>220</sup>	Not a RCT
Toyoda, 2008 <sup>221</sup>	Not a RCT
Tsai, 2008 <sup>222</sup>	Not a RCT
Vivarelli, 2004 <sup>223</sup>	Not a RCT
Wang, 2014 <sup>224</sup>	Duplicate report
Wang, 2015 <sup>225</sup>	No relevant outcome assessed
Wu, 2016 <sup>226</sup>	Not a RCT
Xu, 2012 <sup>227</sup>	Not early HCC patients ( $\leq$ 3 cm tumour)
Xu, 2013 <sup>228</sup>	Duplicate report
Xu, 2015a <sup>229</sup>	Not a RCT
Xu, 2015b <sup>230</sup>	Not early HCC patients ( $\leq$ 3 cm tumour)
Yamamoto, 2001 <sup>231</sup>	Not a RCT
Yamasaki, 2011 <sup>232</sup>	Not early HCC patients ( $\leq$ 3 cm tumour)
Yin, 2014 <sup>233</sup>	Not early HCC patients ( $\leq$ 3 cm tumour)
Yin, 2015 <sup>234</sup>	Duplicate report
Yin, 2019 <sup>235</sup>	Not a RCT
Yu, 2016 <sup>236</sup>	Duplicate report
Yuan, 2017 <sup>237</sup>	Not early HCC patients ( $\leq$ 3 cm tumour)
Yuen, 2003 <sup>238</sup>	Not early HCC patients ( $\leq$ 3 cm tumour)
Yun, 2011 <sup>239</sup>	Not a RCT
Zeng, 2018 <sup>240</sup>	Not early HCC patients ( $\leq$ 3 cm tumour)
Zhang, 2013 <sup>241</sup>	Not a RCT
Zhang, 2015 <sup>242</sup>	Not a RCT
Zhang, 2016 <sup>243</sup>	Not a RCT
Zhao, 2011 <sup>244</sup>	Not early HCC patients ( $\leq$ 3 cm tumour)
Zhi, 2006 <sup>245</sup>	Not a RCT
Zhou, 2009 <sup>246</sup>	Not early HCC patients ( $\leq$ 3 cm tumour)
Zhou, 2014 <sup>247</sup>	Not a RCT
Zhou, 2019 <sup>248</sup>	Not early HCC patients ( $\leq$ 3 cm tumour)
Zhu, 2007 <sup>249</sup>	Not early HCC patients ( $\leq$ 3 cm tumour)
Zhu, 2019 <sup>250</sup>	Not early HCC patients ( $\leq$ 3 cm tumour)

## Appendix 3 Risk of bias assessment results (randomised controlled trials)

Trial	ROB arising from the randomisation process	ROB due to deviations from the intended intervention	ROB due to missing outcome data	ROB in measurement of the outcome	ROB in selection of the reported result	Overall judgement of ROB
Abdelaziz, 2014 – Complete response <sup>64</sup>	High	Low	Low	Low	Low	High
Aikata, 2006 – OS (abstract) <sup>43</sup>	Some concerns	High	High	Low	Low	High
Aikata, 2006 – PFS (abstract) <sup>43</sup>	Some concerns	High	High	Low	Some concerns	High
Azab, 2011 – Complete response <sup>65</sup>	Some concerns	Low	Low	Low	Low	Some concerns
Bian, 2014 – Recurrence <sup>78</sup>	Some concerns	Low	Low	Low	Low	Some concerns
Brunello, 2008 – OS <sup>44</sup>	Low	Low	Low	Low	Low	Low
Chen, 2005 – OS <sup>67</sup>	Some concerns	Low	Low	Low	Low	Some concerns
Chen, 2005 – OS <sup>66</sup>	Some concerns	Low	High	Low	Low	High
Chen, 2006 – OS <sup>68</sup>	Low	High	Low	Low	Low	High
Chen, 2006 – PFS <sup>68</sup>	Low	High	Low	Low	Some concerns	High
Chen, 2014 – OS <sup>45</sup>	Low	Low	Low	Low	Low	Low
Chen, 2014 – Recurrence <sup>45</sup>	Low	Low	Low	Low	Low	Low
Fang, 2014 – OS <sup>46</sup>	Some concerns	Low	Low	Low	Low	Some concerns
Fang, 2014 – Disease-free survival <sup>46</sup>	Some concerns	Low	Low	Low	Some concerns	Some concerns
Fang, 2014 – Recurrence <sup>46</sup>	Some concerns	Low	Low	Low	Low	Some concerns
Ferrari, 2007 – OS <sup>77</sup>	Some concerns	Low	Low	Low	Low	Some concerns
Gan, 2004 – Recurrence <sup>47</sup>	Some concerns	Low	High	Low	Low	High
Giorgio, 2011 – OS <sup>48</sup>	Low	High	High	Low	Low	High
Giorgio, 2011 – Recurrence <sup>48</sup>	Low	High	High	Low	Low	High

continued

Trial	ROB arising from the randomisation process	ROB due to deviations from the intended intervention	ROB due to missing outcome data	ROB in measurement of the outcome	ROB in selection of the reported result	Overall judgement of ROB
Huang, 2005 – OS <sup>49</sup>	High	High	High	Low	Low	High
Huang, 2005 – PFS <sup>49</sup>	High	High	High	Low	Low	High
Huang, 2005 – Recurrence <sup>49</sup>	High	High	High	Low	Some concerns	High
Huang, 2010 – OS <sup>69</sup>	Low	Low	High	Low	Low	High
Huo, 2003 – OS <sup>70</sup>	High	Low	High	Low	Low	High
Izumi, 2019 – PFS (abstract) <sup>50</sup>	Some concerns	High	High	Low	Some concerns	High
Kim, 2020 – OS <sup>51</sup>	Some concerns	Low	Low	Low	Low	Some concerns
Kim, 2020 – PFS <sup>51</sup>	Some concerns	Low	Low	Low	Low	Some concerns
Koda, 2001 – OS <sup>52</sup>	Low	Low	Low	Low	Low	Low
Koda, 2001 – Recurrence <sup>52</sup>	Low	Low	Low	Low	Low	Low
Lencioni, 2003 – OS <sup>71</sup>	Low	Low	Low	Low	Low	Low
Lencioni, 2003 – PFS <sup>71</sup>	Low	Low	Low	Low	Low	Low
Lencioni, 2003 – Recurrence <sup>71</sup>	Low	Low	Low	Low	Low	Low
Lin, 2004 – OS <sup>59</sup>	Low	Low	Low	Low	Low	Low
Lin, 2004 – PFS <sup>59</sup>	Low	Low	Low	Low	Low	Low
Lin, 2004 – Recurrence <sup>59</sup>	Low	Low	Low	Low	Low	Low
Lin, 2005 – OS <sup>53</sup>	Some concerns	Low	Low	Low	Low	Some concerns
Lin, 2005 – PFS <sup>53</sup>	Some concerns	Low	Low	Low	Low	Some concerns
Liu, 2016 – OS <sup>72</sup>	Some concerns	Low	Low	Low	Low	Some concerns
Mizuki, 2010 – OS <sup>60</sup>	Low	Some concerns	Some concerns	Low	Low	Some concerns
Mizuki, 2010 – PFS <sup>60</sup>	Low	Some concerns	Some concerns	Low	Low	Some concerns
Ng, 2017 – OS <sup>79</sup>	Low	Low	Low	Low	Low	Low
Ng, 2017 – PFS <sup>79</sup>	Low	Low	Low	Low	Low	Low
Orlacchio, 2014 – OS <sup>61</sup>	Some concerns	Low	Some concerns	Low	Low	Some concerns
Peng, 2012 – OS <sup>74</sup>	Low	Low	Low	Low	Low	Low

Trial	ROB arising from the randomisation process	ROB due to deviations from the intended intervention	ROB due to missing outcome data	ROB in measurement of the outcome	ROB in selection of the reported result	Overall judgement of ROB
Peng, 2012 – Recurrence-free survival <sup>74</sup>	Low	Low	Low	Low	Low	Low
Shibata, 2002 – Recurrence <sup>62</sup>	Low	Low	Some concerns	Low	Low	Some concerns
Shibata, 2009 – OS <sup>54</sup>	High	Low	Low	Low	Low	High
Shibata, 2009 – PFS <sup>54</sup>	High	Low	Low	Low	Low	High
Shiina, 2005 – OS <sup>55</sup>	Some concerns	Low	Low	Low	Low	Some concerns
Vietti Violi, 2018 – OS <sup>63</sup>	Low	Some concerns	Low	Low	Low	Some concerns
Vietti Violi, 2018 – Progression <sup>63</sup>	Low	Some concerns	Low	Low	Low	Some concerns
Vietti Violi, 2018 – TTP <sup>63</sup>	Low	Some concerns	Low	Low	Low	Some concerns
Xia, 2020 – OS <sup>75</sup>	Low	Low	Low	Low	Low	Low
Xia, 2020 – Recurrence-free survival <sup>75</sup>	Low	Low	Low	Low	Low	Low
Yan, 2016 – OS <sup>56</sup>	High	Some concerns	Low	Low	Low	High
Yan, 2016 – PFS <sup>56</sup>	High	Some concerns	Low	Low	Low	High
Zhang, 2007 – OS <sup>76</sup>	Low	Low	Low	Low	Low	Low
Zou, 2017 – Response <sup>57</sup>	Some concerns	Low	Low	Low	Low	Some concerns
<b>Total</b>	<b>High: 9 Some concerns: 20 Low: 29</b>	<b>High: 10 Some concerns: 7 Low: 41</b>	<b>High: 12 Some concerns: 4 Low: 42</b>	<b>Low: 58</b>	<b>Some concerns: 5 Low: 53</b>	<b>High: 19 Some concerns: 21 Low: 18</b>



## **Appendix 4** Characteristics and results of randomised controlled trials included in the review

Study name and location	Participant information	Intervention	Comparator	Main results	ROB
<b>RFA vs. MWA</b>					
Abdelaziz, 2014 <sup>64</sup> Egypt	111 patients (with 128 tumours) ≤ 5 cm; subgroup of 87 tumours ≤ 3 cm	RFA (n = 32 tumours ≤ 3 cm)	MWA (n = 55 tumours ≤ 3 cm)	<p><b>Response:</b> RFA: Complete ablation: 30/32 (93.8%) tumours; partial ablation: 2/32 (6.2%) tumours. MWA: Complete ablation: 54/55 (98.2%) tumours; partial ablation: 1/55 (1.8%) tumours.</p> <p><b>Adverse events:</b> There were no major complications or deaths in either group. No other results reported for ≤ 3 cm tumour subgroup.</p>	High
Shibata, 2002 <sup>62</sup> Japan	72 patients (with 94 tumours) < 4 cm; subgroup of 88 tumours ≤ 3 cm	RFA (n = 36 patients, 3 of whom had tumours 3–4 cm)	MWA (n = 36 patients, 3 of whom had tumours 3–4 cm)	<p><b>Response:</b> RFA: 46/48 (96%) nodules showed complete response and 2 (4%) had residual lesions or incomplete response. MWA: 41/46 (89%) nodules showed complete response and 5 (11%) showed incomplete response.</p> <p><b>Adverse events:</b> RFA: one major complication (3%) (segmental hepatic infarction). MWA: four major complications (11%) (liver abscess, cholangitis, subcutaneous abscess with skin burn and subcapsular hematoma). No other results reported, other than Kaplan–Meier curve.</p>	Some concerns
Viotti Violi, 2018 <sup>63</sup> Switzerland and France	152 patients with tumours ≤ 4 cm (mean tumour size 1.8 cm, <8% patients had tumours > 3 cm)	MWA (n = 76 patients, 6 of whom had tumours 3–4 cm)	RFA (n = 76 patients, 5 of whom had tumours 3–4 cm)	<p><b>OS:</b> MWA: 2 years: 86%. RFA: 2 years: 84%.</p> <p><b>Recurrence:</b> MWA: 2 years: 6% (local tumour progression). RFA: 2 years: 12% (local tumour progression). Comparison between groups: RR 1.62, 95% CI 0.66 to 3.94.</p> <p><b>TTP:</b> MWA: Median 12 months (95% CI 5 to 28). RFA: Median 16 months (95% CI 4 to 24). Comparison between groups: HR 0.72, 95% CI 0.44 to 1.18.</p> <p><b>Response:</b> MWA: 95% achieved complete response after one treatment; 5% achieved complete response after two treatments. RFA: 96% achieved complete response after one treatment; 4% achieved complete response after two treatments.</p> <p><b>Adverse events:</b> MWA: 2 (2%) grade IV AEs; 5 (5%) grade I–II AEs. RFA: 3 (3%) grade III AEs; 12 (11.5%) grade I–II AEs. No treatment-related deaths in either group.</p>	Some concerns

Study name and location	Participant information	Intervention	Comparator	Main results	ROB
<b>RFA vs. PEI</b>					
Azab, 2011 <sup>65</sup> Egypt	90 patients (with 98 tumours) ≤ 5 cm; subgroup of 48 tumours ≤ 3 cm	PEI + RFA (n = 16 tumours ≤ 3 cm)	RFA alone (n = 16 tumours ≤ 3 cm) PEI alone (n = 16 tumours ≤ 3 cm)	<p><b>Response:</b> PEI + RFA: 15/16 (93.8%) nodules had complete ablation and 1/16 (6.2%) had partial ablation after one session; 16/16 (100%) nodules had complete ablation after two sessions. RFA alone: 12/16 (75%) nodules had complete ablation and 4/16 (25%) had partial ablation after one session; 14/16 (87.5%) nodules had complete ablation and 2/16 (12.5%) had partial ablation after two sessions. PEI alone: 0% nodules had complete ablation after two sessions; 13/16 (81.25%) nodules had complete ablation and 3/16 (18.75%) had partial ablation after all sessions.</p> <p><b>Adverse events:</b> There were no mortalities related to any of the techniques. No other results reported for ≤ 3cm tumour subgroup.</p>	Some concerns
Brunello, 2008 <sup>44</sup> Italy	139 patients with tumours ≤ 3 cm	RFA (n = 70 patients)	PEI (n = 69 patients)	<p><b>OS:</b> RFA: Number of events (death): 26. PEI: Number of events (death): 28. Comparison between groups: HR 0.82, 95% CI 0.48 to 1.41. Adjusted HR 0.88, 95% CI 0.50 to 1.53.</p> <p><b>Distant intrahepatic recurrence:</b> RFA: Number of events: 32. PEI: Number of events: 35.</p> <p><b>Response:</b> RFA: 1-year response: 46/70 (65.7%). Early complete response (30-50 days): 67/70 (95.7%). PEI: 1-year response: 25/69 (36.2%). Early complete response (30-50 days, patients with 1-year follow-up only): 42/64 (65.6%).</p> <p><b>Adverse events:</b> RFA: Treatment-emergent AEs: 10 (14.3%) patients; major complications: 2 patients. PEI: Treatment-emergent AEs: 12 (17.4%) patients; major complications: 2 patients.</p> <p><b>Economic outcomes:</b> Mean direct medical costs were €4097 for PEI and €6540 for RFA. ICER for using RFA instead of PEI was €8286 (95% CI €2742 to €20,917).</p>	Low

continued

Study name and location	Participant information	Intervention	Comparator	Main results	ROB
Giorgio, 2011 <sup>48</sup> Italy	285 patients with tumours ≤ 3 cm	RFA (n = 142 patients)	PEI (n = 143 patients)	<p><b>OS:</b> RFA: 1 year: 95%; 2 years: 90%; 3 years: 83%; 4 years: 73%; 5 years: 70%. PEI: 1 year: 95%; 2 years: 83%; 3 years: 78%; 4 years: 70%; 5 years: 68%. Comparison between groups: HR 0.81, 95% CI 0.46 to 1.39.</p> <p><b>Local recurrence:</b> RFA: 1 year: 4.1%; 2 years: 5.7%; 3 years: 7.8%; 4 years: 8.9%; 5 years: 11.7%. PEI: 1 year: 5.2%; 2 years: 6.7%; 3 years: 9.4%; 4 years: 11.5%; 5 years: 12.8%.</p> <p><b>Adverse events:</b> RFA: Major complication: 0.9%. PEI: Major complication: 1.9%. No deaths related to the procedure in either group and no cases of seeding.</p> <p><b>Economic outcomes:</b> In Italy the medical costs of both procedures were the same so cost calculated only on basis of technical material. The overall cost for PEI needles plus ethanol was €1359 plus VAT, the cost of the generator plus the RFA electrode needles was €171,000 plus VAT, with a statistically significant difference (p &lt; 0.001).</p>	High
Lencioni, 2003 <sup>71</sup> Authors from Italy and Germany	104 patients with tumours ≤ 5 cm (large portion had tumours ≤ 3 cm)	PEI (n = 50 patients with mean tumour size 2.8 cm; 84% had tumours ≤ 3 cm)	RFA (n = 54 patients with mean tumour size 2.8 cm; 88% had tumours ≤ 3 cm)	<p><b>OS:</b> PEI: 1 year: 96%; 2 years: 88%. RFA: 1 year: 100%; 2 years: 98%. PEI: Number of events (death): 5. RFA: Number of events (death): 1.</p> <p><b>Event-free survival:</b> PEI: 1 year: 77%; 2 years: 43%. RFA: 1 year: 86%; 2 years: 64%.</p> <p><b>Local recurrence:</b> PEI: Number of events: 13. RFA: Number of events: 3.</p> <p><b>Response:</b> PEI: 60/73 (82%) tumours had complete response after one cycle; 69/73 (94.5%) tumours had complete response after two cycles. RFA: 63/69 (91%) tumours had complete response after one session; 68/69 (98.6%) tumours had complete response after two sessions.</p> <p><b>Adverse events:</b> PEI: No procedure-related death, haemorrhage, infection, needle-track seeding or hepatic failure. Mild to moderate pain requiring analgesics: 13; fever: 5; chemical thrombosis of a portal venous branch: 1. RFA: No procedure-related death, haemorrhage, infection, needle-track seeding or hepatic failure. Mild to moderate pain requiring analgesics: 15; fever: 10; pleural effusions: 4 lesions; asymptomatic arteriovenous shunts: 3.</p>	Low

Study name and location	Participant information	Intervention	Comparator	Main results	ROB
Lin, 2004 <sup>59</sup> Authors from Taiwan	157 patients with tumours ≤ 4 cm; subgroup of 114 patients with tumours ≤ 3 cm	RFA (n = 37 patients with tumours ≤ 3 cm)	Low-dose PEI (n = 38 patients with tumours ≤ 3 cm) High-dose PEI (n = 39 patients with tumours ≤ 3 cm)	<p><b>OS:</b></p> <p>RFA in patients with tumours 1–2 cm: 1 year: 96%; 2 years: 84%; 3 years: 78%. RFA in patients with tumours 2.1–3 cm: 1 year: 89%; 2 years: 78%; 3 years: 73%. Low-dose PEI in patients with tumours 1–2 cm: 1 year: 94%; 2 years: 78%; 3 years: 70%. Low-dose PEI in patients with tumours 2.1–3 cm: 1 year: 84%; 2 years: 70%; 3 years: 62%. High-dose PEI in patients with tumours 1–2 cm: 1 year: 93%; 2 years: 80%; 3 years: 72%. High-dose PEI in patients with tumours 2.1–3 cm: 1 year: 83%; 2 years: 71%; 3 years: 64%.</p> <p><b>Cancer-free survival:</b></p> <p>RFA in patients with tumours 1–2 cm: 1 year: 84%; 2 years: 64%; 3 years: 49%. RFA in patients with tumours 2.1–3 cm: 1 year: 76%; 2 years: 53%; 3 years: 40%. Low-dose PEI in patients with tumours 1–2 cm: 1 year: 73%; 2 years: 57%; 3 years: 41%. Low-dose PEI in patients with tumours 2.1–3 cm: 1 year: 68%; 2 years: 57%; 3 years: 30%. High-dose PEI in patients with tumours 1–2 cm: 1 year: 71%; 2 years: 61%; 3 years: 43%. High-dose PEI in patients with tumours 2.1–3 cm: 1 year: 63%; 2 years: 51%; 3 years: 32%.</p> <p><b>Local tumour progression:</b></p> <p>RFA in patients with tumours 1–2 cm: 1 year: 4%; 2 years: 9%; 3 years: 9%. RFA in patients with tumours 2.1–3 cm: 1 year: 11%; 2 years: 18%; 3 years: 18%. Low-dose PEI in patients with tumours 1–2 cm: 1 year: 8%; 2 years: 13%; 3 years: 13%. Low-dose PEI in patients with tumours 2.1–3 cm: 1 year: 18%; 2 years: 37%; 3 years: 37%. High-dose PEI in patients with tumours 1–2 cm: 1 year: 7%; 2 years: 12%; 3 years: 12%. High-dose PEI in patients with tumours 2.1–3 cm: 1 year: 21%; 2 years: 33%; 3 years: 33%.</p> <p><b>Adverse events:</b></p> <p>RFA: One patient (1.9%) had transient pleural effusion, although not reported whether their tumour was &lt; 3cm or not. No other severe adverse effect was observed. Low-dose PEI: No severe adverse effect was observed. High-dose PEI: No severe adverse effect was observed.</p> <p><b>Economic outcomes:</b></p> <p>RFA: Mean hospital stay: 4.4 days (range 3–15). Low-dose PEI: Mean hospital stay: 1.6 days (range 2–3). High-dose PEI: Mean hospital stay: 2.1 days (range 2–4).</p>	Low

continued

Study name and location	Participant information	Intervention	Comparator	Main results	ROB
Lin, 2005 <sup>53</sup> Taiwan	187 patients with tumours ≤ 3 cm	RFA (n = 62 patients)	PEI (n = 62 patients) PAI (n = 63 patients)	<p><b>OS:</b> RFA: 1 year: 93%; 2 years: 81%; 3 years: 74%. PEI: 1 year: 88%; 2 years: 66%; 3 years: 51%. PAI: 1 year: 90%; 2 years: 67%; 3 years: 53%. RFA: Number of events (death): 10 (16.1%). PEI: Number of events (death): 17 (27.4%). PAI: Number of events (death): 15 (23.8%).</p> <p><b>Cancer-free survival:</b> RFA: 1 year: 74%; 2 years: 60%; 3 years: 43%. PEI: 1 year: 70%; 2 years: 41%; 3 years: 21%. PAI: 1 year: 71%; 2 years: 43%; 3 years: 23%.</p> <p><b>Recurrence:</b> RFA: 1 year: 10%; 2 years: 14%; 3 years: 14%. PEI: 1 year: 16%; 2 years: 34%; 3 years: 34%. PAI: 1 year: 14%; 2 years: 31%; 3 years: 31%. RFA: Number of local recurrence events: 8/60 (13.3%). PEI: Number of local recurrence events: 19/55 (34.5%). PAI: Number of local recurrence events: 17/58 (29.3%).</p> <p><b>Complete response:</b> RFA: 96.1% (75/78 tumours). PEI: 88.1% (67/76 tumours). PAI: 92.4% (73/79 tumours).</p> <p><b>Adverse events:</b> RFA: 4.8% (3/62 patients) had serious AEs (2 patients with haemothorax and 1 with gastric bleeding and perforation). PEI: No serious AEs. PAI: No serious AEs.</p> <p><b>Economic outcomes:</b> RFA: Mean hospitalisation: 4.2 days (range 3–18). PEI: Mean hospitalisation: 1.7 days (range 2–3). PAI: Mean hospitalisation: 2.2 days (range 2–5).</p>	Some concerns

Study name and location	Participant information	Intervention	Comparator	Main results	ROB
Shiina, 2005 <sup>55</sup> Japan	232 patients with tumours $\leq 3$ cm	RFA (n = 118 patients)	PEI (n = 114 patients)	<p><b>OS:</b>  RFA: 4 years: 74% (95% CI 65% to 84%).  PEI: 4 years: 57% (95% CI 45% to 71%).  RFA: Number of events (death): 25 (21.2%).  PEI: Number of events (death): 40 (35.1%).  Comparison between groups: RR 0.54, 95% CI 0.33 to 0.89.</p> <p><b>Recurrence:</b>  RFA: Number of events: 78 (66.1%) (new lesion: 74; local tumour progression: 2; extrahepatic recurrence: 2).  PEI: Number of events: 90 (78.9%) (new lesion: 73; local tumour progression: 13; extrahepatic recurrence: 4).  Comparison between groups: RR 0.57, 95% CI 0.41 to 0.80.  Local tumour progression: RR 0.12, 95% CI 0.03 to 0.55.</p> <p><b>Adverse events:</b>  RFA: AEs grade <math>\geq 3</math>: 6 (5.1%) (1 transient jaundice, 1 skin burn, 1 hepatic infarction and 3 seeding of malignant cells).  PEI: AEs grade <math>\geq 3</math>: 3 (2.6%) (1 liver abscess and 2 neoplastic seeding).</p> <p><b>Economic outcomes:</b>  RFA: Length of hospitalisation: 10.8 <math>\pm</math> 5.5 days.  PEI: Length of hospitalisation: 26.1 <math>\pm</math> 9.9 days.</p>	Some concerns
					continued

Study name and location	Participant information	Intervention	Comparator	Main results	ROB
<b>RFA vs. PAI</b>					
Lin, 2005 <sup>53</sup> Taiwan	187 patients with tumours ≤3cm	RFA (n = 62 patients)	PEI (n = 62 patients) PAI (n = 63 patients)	<p><b>OS:</b> RFA: 1 year: 93%; 2 years: 81%; 3 years: 74%. PEI: 1 year: 88%; 2 years: 66%; 3 years: 51%. PAI: 1 year: 90%; 2 years: 67%; 3 years: 53%. RFA: Number of events (death): 10 (16.1%). PEI: Number of events (death): 17 (27.4%). PAI: Number of events (death): 15 (23.8%).</p> <p><b>Cancer-free survival:</b> RFA: 1 year: 74%; 2 years: 60%; 3 years: 43%. PEI: 1 year: 70%; 2 years: 41%; 3 years: 21%. PAI: 1 year: 71%; 2 years: 43%; 3 years: 23%.</p> <p><b>Recurrence:</b> RFA: 1 year: 10%; 2 years: 14%; 3 years: 14%. PEI: 1 year: 16%; 2 years: 34%; 3 years: 34%. PAI: 1 year: 14%; 2 years: 31%; 3 years: 31%. RFA: Number of local recurrence events: 8/60 (13.3%). PEI: Number of local recurrence events: 19/55 (34.5%). PAI: Number of local recurrence events: 17/58 (29.3%).</p> <p><b>Complete response:</b> RFA: 96.1% (75/78 tumours). PEI: 88.1% (67/76 tumours). PAI: 92.4% (73/79 tumours).</p> <p><b>Adverse events:</b> RFA: 4.8% (3/62 patients) had serious AEs (2 patients with haemothorax and 1 with gastric bleeding and perforation). PEI: No serious AEs. PAI: No serious AEs.</p> <p><b>Economic outcomes:</b> RFA: Mean hospitalisation: 4.2 days (range 3–18). PEI: Mean hospitalisation: 1.7 days (range 2–3). PAI: Mean hospitalisation: 2.2 days (range 2–5).</p>	Some concerns
<b>RFA vs. laser ablation</b>					
Ferrari, 2007 <sup>77</sup> Authors from Italy	81 patients with tumours ≤ 4 cm; subgroup of 28 patients with tumours ≤ 2.5 cm	Laser ablation (n = 12 patients with tumours ≤ 2.5 cm)	RFA (n = 16 patients with tumours ≤ 2.5 cm)	<p><b>Adverse events:</b> No deaths or major or minor complications occurred during the procedures in either group and no cases of neoplastic seeding observed. No other results reported for ≤ 2.5 cm subgroup, other than Kaplan–Meier curve.</p>	Some concerns

Study name and location	Participant information	Intervention	Comparator	Main results	ROB
Orlacchio, 2014 <sup>61</sup> Italy	30 patients with tumours ≤ 4 cm (mean tumour size 2.4 cm)	Laser ablation (n = 15 patients with mean tumour size 2.34 cm)	RFA (n = 15 patients with mean tumour size 2.41 cm)	<p><b>OS:</b> Laser ablation: Number of events (death): 0 RFA: Number of events (death): 0</p> <p><b>PFS:</b> Laser ablation: 3 months: 85%; 6 months: 62%; 1 year: 54%. RFA: 3 months: 92%; 6 months: 86%; 1 year: 86%.</p> <p><b>Recurrence:</b> Laser ablation: Number of events (local disease progression): 6 (40%). RFA: Number of events (local disease progression): 2 (13.3%).</p> <p><b>Complete response:</b> Laser ablation: 66.7% nodules (10/15) after first procedure; 3/5 patients after a second procedure (87% in total). RFA: 86.7% nodules (13/15) after first procedure; 1/2 patients after a second procedure (93% in total).</p> <p><b>Adverse events:</b> Laser ablation: 2 (13.3%) (1 vasovagal reaction and 1 postablation syndrome). RFA: 14 (93.3%) (2 vasovagal reactions and 12 postablation syndrome). No major complications in either arm.</p>	Some concerns
Zou, 2017 <sup>57</sup> China	74 patients with tumours ≤ 3 cm	Laser ablation (n = 35 patients)	RFA (n = 39 patients)	<p><b>Response:</b> Laser ablation: 31 (88.6%) had a complete response; 4 (11.4%) had a partial response. RFA: 36 (92.3%) had a complete response; 3 (7.7%) had a partial response.</p> <p><b>Patient satisfaction:</b> Self-made satisfaction questionnaire including intraoperative discomfort, postoperative therapy effects, adverse reactions and physical recovery. Maximum score of 100, 81–100 = great satisfaction, 61–80 = general satisfaction, &lt; 60 = dissatisfaction. Laser ablation: great satisfaction: 30 (85.7%), general satisfaction: 3 (8.6%), dissatisfaction: 2 (5.7%). RFA: great satisfaction: 25 (64.1%), general satisfaction: 2 (5.1%), dissatisfaction: 12 (30.8%).</p> <p><b>Adverse events:</b> Laser ablation: Fever: 4 (11.4%); nausea: 17 (48.6%); vomiting: 10 (28.6%); diarrhoea: 1 (2.9%); abdominal pain: 26 (74.3%); skin rash: 1 (2.9%). RFA: Fever: 5 (12.8%); nausea: 19 (48.7%); vomiting: 12 (30.8%); diarrhoea: 1 (2.6%); abdominal pain: 29 (74.4%); skin rash: 3 (7.7%).</p>	Some concerns

continued

Study name and location	Participant information	Intervention	Comparator	Main results	ROB
<b>RFA vs. resection</b>					
Chen, 2005 (reported in Chinese) <sup>67</sup> China	132 patients with tumours ≤ 5 cm; subgroup of 55 patients with tumours ≤ 3 cm	Resection (n = 31 patients with tumours ≤ 3 cm)	RFA (n = 24 patients with tumours ≤ 3 cm)	No results reported for ≤ 3 cm subgroup, other than Kaplan–Meier curve.	Some concerns
Fang, 2014 <sup>46</sup> China	120 patients with tumours ≤ 3 cm	RFA (n = 60 patients)	Hepatectomy (n = 60 patients)	<p><b>OS:</b> RFA: 1 year: 97.5%; 2 years: 91.2%; 3 years: 82.5%. Hepatectomy: 1 year: 93.7%; 2 years: 86.2%; 3 years: 77.5%.</p> <p><b>Disease-free survival:</b> RFA: 1 year: 91.6%; 2 years: 87.4%; 3 years: 55.4%. Hepatectomy: 1 year: 90.4%; 2 years: 85.2%; 3 years: 41.3%.</p> <p><b>Recurrence:</b> RFA: Number of events: 22. Hepatectomy: Number of events: 21.</p> <p><b>Response:</b> RFA: Complete tumour treatment: 57/60 (95.0%). Hepatectomy: Complete tumour treatment: 58/60 (96.7%).</p> <p><b>Adverse events:</b> RFA: Postoperative complications: 2. Major complications: 1. Hospital mortality: 0. Serious pain requiring analgesic (not included in postoperative complications): 3. Hepatectomy: Postoperative complications: 17. Major complications: 14. Hospital mortality: 0. Serious pain requiring analgesic (not included in postoperative complications): 43.</p> <p><b>Economic outcomes:</b> RFA: ICU stay: 0. Hospital days (mean): 4.3 (SD 1.5). Hepatectomy: ICU stay: 6 (10%). Hospital days (mean): 11.8 (SD 3.1).</p>	Some concerns
Feng, 2012 <sup>58</sup> China	168 patients with tumours < 4 cm; subgroup of 56 patients with tumours ≤ 2 cm	RFA (n = 31 patients with tumours ≤ 2 cm)	Surgical resection (n = 25 patients with tumours ≤ 2 cm)	<p>No results reported for ≤ 2 cm subgroup.</p> <p><b>Adverse events:</b> No treatment-related mortality.</p>	N/A

Study name and location	Participant information	Intervention	Comparator	Main results	ROB
Huang, 2010 <sup>69</sup> China	230 patients with tumours ≤ 5 cm; subgroup of 159 patients with tumours ≤ 3 cm	RFA (n = 88 patients with tumours ≤ 3 cm)	Resection (n = 71 patients with tumours ≤ 3 cm)	<p><b>OS:</b> RFA in patients with solitary tumours ≤ 3 cm (n = 57): 1 year: 91.2%; 2 years: 84.2%; 3 years: 77.2%; 4 years: 71.9%; 5 years: 61.4%. RFA in patients with multifocal tumours ≤ 3 cm (n = 31): 1 year: 77.4%; 2 years: 64.5%; 3 years: 58.1%; 4 years: 58.1%; 5 years: 45.2%. Resection in patients with solitary tumours ≤ 3 cm (n = 45): 1 year: 100%; 2 years: 97.8%; 3 years: 95.6%; 4 years: 86.7%; 5 years: 82.2%. Resection in patients with multifocal tumours ≤ 3 cm (n = 26): 1 year: 92.3%; 2 years: 88.5%; 3 years: 80.8%; 4 years: 73.1%; 5 years: 69.2%.</p> <p><b>Adverse events:</b> No deaths within same hospital admission. No other results reported for ≤ 3 cm subgroup, other than separate Kaplan–Meier curves for solitary tumours ≤ 3 cm and multifocal tumours ≤ 3 cm.</p> <p><b>Recurrence-free survival:</b> RFA: 3 years: 47.7%. Surgery: 3 years: 49.8%. Comparison between groups: HR 0.96.</p> <p><b>Adverse events:</b> No perioperative mortality in either group.</p> <p><b>OS:</b> Resection: 1 year: 100%; 3 years: 93%; 5 years: 76%. RFA: 1 year: 100%; 3 years: 89%; 5 years: 69%.</p> <p><b>Disease-free survival:</b> Resection: 1 year: 83%; 3 years: 66%; 5 years: 52%. RFA: 1 year: 77%; 3 years: 62%; 5 years: 46%.</p> <p><b>Economic outcomes:</b> Resection: Median hospital stay: 7 days. RFA: Median hospital stay: 4 days.</p> <p><b>OS:</b> Comparison between group: In HR 0.05, 95% CI -0.4 to 0.5. HR 1.05, 95% CI 0.67 to 1.65 (calculated by CRD).</p> <p><b>Repeat-recurrence-free survival:</b> Comparison between group: In HR 0.07, 95% CI -0.34 to 0.47. HR 1.07, 95% CI 0.71 to 1.6 (calculated by CRD).</p>	High
Izumi, 2019 <sup>80</sup> (conference abstract) Japan	308 patients with tumours ≤ 3 cm	RFA (n = 148 analysed)	Surgery (n = 145 analysed)	<p><b>Recurrence-free survival:</b> RFA: 3 years: 47.7%. Surgery: 3 years: 49.8%. Comparison between groups: HR 0.96.</p> <p><b>Adverse events:</b> No perioperative mortality in either group.</p>	High
Ng, 2017 <sup>73</sup> China	218 patients with tumours ≤ 5 cm; subgroup of 55 patients with tumours ≤ 2 cm	Resection (n = 29 patients with tumours ≤ 2 cm)	RFA (n = 26 patients with tumours ≤ 2 cm)	<p><b>OS:</b> Resection: 1 year: 100%; 3 years: 93%; 5 years: 76%. RFA: 1 year: 100%; 3 years: 89%; 5 years: 69%.</p> <p><b>Disease-free survival:</b> Resection: 1 year: 83%; 3 years: 66%; 5 years: 52%. RFA: 1 year: 77%; 3 years: 62%; 5 years: 46%.</p> <p><b>Economic outcomes:</b> Resection: Median hospital stay: 7 days. RFA: Median hospital stay: 4 days.</p>	Low
Xia, 2020 <sup>75</sup> China	240 patients with recurrent HCC tumours ≤ 5 cm; subgroup of 159 patients with tumours ≤ 3 cm	RFA (78 patients with tumours ≤ 3 cm)	Repeat hepatectomy (81 patients with tumours ≤ 3 cm)	<p><b>OS:</b> Comparison between group: In HR 0.05, 95% CI -0.4 to 0.5. HR 1.05, 95% CI 0.67 to 1.65 (calculated by CRD).</p> <p><b>Repeat-recurrence-free survival:</b> Comparison between group: In HR 0.07, 95% CI -0.34 to 0.47. HR 1.07, 95% CI 0.71 to 1.6 (calculated by CRD).</p>	Low
Zhu, 2021 (protocol) <sup>80</sup> China	Ongoing RCT	RFA	Laparoscopic hepatectomy	N/A	N/A

continued

Study name and location	Participant information	Intervention	Comparator	Main results	ROB
<b>RFA vs. proton beam radiotherapy</b>					
Kim, 2021 <sup>162</sup> South Korea	144 patients with recurrent/residual tumours < 3 cm	Proton beam radiotherapy (n = 72 patients)	RFA (n = 72 patients)	<p><b>OS:</b> Proton beam radiotherapy: 2 years: 91.7%; 3 years: 80.8%; 4 years: 75.4%. RFA: 2 years: 90.3%; 3 years: 86.0%; 4 years: 77.0%. Comparison between groups: HR (2 years) 1.07, 95% CI 0.58 to 1.98.</p> <p><b>PFS:</b> Proton beam radiotherapy: Median 13.4 months (90% CI 7.69 to 16.76 months). RFA: Median 13.7 months (90% CI 9.86 to 18.89 months). Proton beam radiotherapy: 2 years: 31.9%; 3 years: 26.3%; 4 years: 18.7%. RFA: 2 years: 31.9%; 3 years: 17.9%; 4 years: 12.6%. HR (2 years) 0.99, 95% CI 0.70 to 1.41. Proton beam radiotherapy: Number of events (progression): 56 RFA: Number of events (progression): 62</p> <p><b>Adverse events:</b> Proton beam radiotherapy: AEs grade ≥ 3: 0. RFA: AEs grade ≥ III: 9 (16%). No treatment-related late hepatic failure and death without evidence of disease progression and/or subsequent treatment.</p>	Some concerns
<b>RFA vs. RFA + TACE</b>					
Aikata, 2006 (conference abstract) <sup>143</sup> Authors from Japan	44 patients with tumours < 3 cm	RFA + TACE (n = 21 patients)	RFA alone (n = 23 patients)	<p><b>OS:</b> RFA + TACE: 1 year: 95.2%; 2 years: 95.2%; 3 years: 84%. RFA alone: 1 year: 100%; 2 years: 82.6%; 3 years: 73.9%.</p> <p><b>Local tumour progression:</b> RFA + TACE: 1 year: 9.5%; 2 years: 9.5%; 3 years: 9.5%. RFA alone: 1 year: 4.3%; 2 years: 8.7%; 3 years: 8.7%.</p> <p><b>Adverse events:</b> There were no major complications in either group.</p>	High
Peng, 2012 <sup>24</sup> China	139 patients with recurrent HCC tumours ≤ 5 cm; subgroup of 87 patients with tumours ≤ 3 cm	RFA + TACE (n = 41 patients with tumours ≤ 3 cm)	RFA alone (n = 46 patients with tumours ≤ 3 cm)	<p><b>OS:</b> RFA + TACE: 1 year: 98%; 3 years: 70%; 5 years: 50%. RFA alone: 1 year: 83%; 3 years: 60%; 5 years: 50%.</p> <p><b>Recurrence-free survival:</b> RFA + TACE: 1 year: 90%; 3 years: 48%; 5 years: 48%. RFA alone: 1 year: 86%; 3 years: 26%; 5 years: 26%.</p>	Low

Study name and location	Participant information	Intervention	Comparator	Main results	ROB
Shibata, 2009 <sup>54</sup> Japan	89 patients with tumours ≤ 3 cm	RFA + TACE (n = 46 patients)	RFA alone (n = 43 patients)	<p><b>OS:</b> RFA + TACE: 1 year: 100%; 2 years: 100%; 3 years: 84.8%; 4 years: 72.7%. RFA alone: 1 year: 100%; 2 years: 88.8%; 3 years: 84.5%; 4 years: 74%. RFA + TACE: Number of events (death): 5 (10.9%). RFA alone: Number of events (death): 6 (13.9%).</p> <p><b>Event-free survival:</b> RFA + TACE: 1 year: 71.3%; 2 years: 59.9%; 3 years: 48.8%; 4 years: 36.6%. RFA alone: 1 year: 74.3%; 2 years: 52.4%; 3 years: 29.7%; 4 years: 29.7%.</p> <p><b>Local PFS</b> RFA + TACE: 1 year: 84.6%; 2 years: 81.1%; 3 years: 69.7%; 4 years: 55.8%. RFA alone: 1 year: 88.4%; 2 years: 74.1%; 3 years: 74.1%; 4 years: 61.7%.</p> <p><b>Local tumour progression:</b> RFA + TACE: 1 year: 14.4%; 2 years: 17.6%; 3 years: 17.6%; 4 years: 17.6%. RFA alone: 1 year: 11.4%; 2 years: 14.4%; 3 years: 14.4%; 4 years: 14.4%. RFA + TACE: Number of recurrence events: 8 (17.4%). RFA alone: Number of recurrence events: 6 (13.9%).</p> <p><b>Complete response:</b> RFA + TACE: 100%. RFA alone: 100%.</p> <p><b>Adverse events:</b> RFA + TACE: 1 major complication (2%) (segmental hepatic infarction). RFA alone: 1 major complication (2%) (pseudaneurysm of the anterosuperior branch of the right hepatic artery).</p>	High
<b>RFA vs. RFA + PEI</b>					
Azab, 2011 <sup>65</sup> Egypt	90 patients (with 98 tumours) ≤ 5 cm; subgroup of 48 tumours ≤ 3 cm	PEI + RFA (n = 16 tumours ≤ 3 cm)	RFA alone (n = 16 tumours ≤ 3 cm) PEI alone (n = 16 tumours ≤ 3 cm)	<p><b>Response:</b> PEI + RFA: 15/16 (93.8%) nodules had complete ablation and 1/16 (6.2%) had partial ablation after one session; 16/16 (100%) nodules had complete ablation after two sessions. RFA alone: 12/16 (75%) nodules had complete ablation and 4/16 (25%) had partial ablation after one session; 14/16 (87.5%) nodules had complete ablation and 2/16 (12.5%) had partial ablation after two sessions. PEI alone: 0% nodules had complete ablation after two sessions; 13/16 (81.25%) nodules had complete ablation and 3/16 (18.75%) had partial ablation after all sessions.</p> <p><b>Adverse events:</b> There were no mortalities related to any of the techniques. No other results reported for ≤ 3 cm tumour subgroup.</p>	Some concerns

continued

Study name and location	Participant information	Intervention	Comparator	Main results	ROB
Chen, 2005 (reported in Chinese) <sup>66</sup> China	86 patients with tumours ≤ 5 cm; subgroup of 47 patients with tumours ≤ 3 cm	RFA + PEI (n = 24 patients with tumours ≤ 3 cm)	RFA alone (n = 23 patients with tumours ≤ 3 cm)	<p><b>OS:</b> RFA + PEI: 1 year: 86.8%; 2 years: 79.0%. RFA alone: 1 year: 81.2%; 2 years: 64.9%.</p> <p><b>Recurrence:</b> RFA + PEI: 1 year: 4.3%; 2 years: 20.9%. RFA alone: 1 year: 16.8%; 2 years: 34.1%.</p> <p><b>Adverse events:</b> There were no serious AEs in either group.</p>	High
Zhang, 2007 <sup>70</sup> China	133 patients with tumours ≤ 7 cm; subgroup of 60 patients with tumours ≤ 3 cm	RFA + PEI (29 patients with tumours ≤ 3 cm)	RFA alone (31 patients with tumours ≤ 3 cm)	<p><b>OS:</b> RFA + PEI: 1 year: 96.5%; 2 years: 92.9%; 3 years: 83.6%; 4 years: 77.4%; 5 years: 55.3%. RFA alone: 1 year: 93.5%; 2 years: 80.7%; 3 years: 76.6%; 4 years: 70.2%; 5 years: 50.2%.</p> <p><b>Response:</b> RFA + PEI: 52/66 (78.8%) had complete tumour ablation after one treatment, 14 (21.2%) had complete tumour ablation after a second treatment. RFA alone: 48/67 (71.6%) had complete tumour ablation after one treatment, 18 (26.9%) had complete tumour ablation after a second treatment and 1 patient (1.5%) had viable tumour cells after two treatment sessions. Tumour diameter was a significant prognostic factor for overall recurrence, intrahepatic recurrence and local recurrence.</p>	Low
<b>RFA vs. RFA + iodine-131 metuximab</b>					
Bian, 2014 <sup>78</sup> China	127 patients with BCLC stage 0–B; subgroup of 78 patients with tumours < 3 cm	RFA + iodine-131 metuximab (n = 38 patients with tumours < 3 cm)	RFA alone (n = 40 patients with tumours < 3 cm)	<p><b>Recurrence:</b> Comparison between groups: HR 0.46, 95% CI 0.21 to 1.01.</p> <p><b>Adverse events:</b> There were no serious AEs or treatment-related deaths in either group. No other results reported for &lt; 3 cm tumour subgroup.</p>	Some concerns

Study name and location	Participant information	Intervention	Comparator	Main results	ROB
<b>RFA vs. RFA + iodine-125</b>					
Chen, 2014 <sup>45</sup> China	136 patients with tumours ≤ 3 cm	RFA + iodine-125 (n = 68 patients)	RFA alone (n = 68 patients)	<p><b>OS:</b> RFA + iodine-125: Mean 95.8 months. RFA alone: Mean 70.8 months. RFA + iodine-125: 1 year: 100%; 2 years: 95.6%; 3 years: 86.7%; 4 years: 73.5%; 5 years: 66.1%. RFA alone: 1 year: 95.6%; 2 years: 85.2%; 3 years: 75.0%; 4 years: 58.8%; 5 years: 47.0%. RFA + iodine-125: Number of events (death): 23. RFA alone: Number of events (death): 36. Comparison between groups: HR 0.502, 95% CI 0.313 to 0.806.</p> <p><b>Recurrence:</b> RFA + iodine-125: 1 year: 4.5%; 2 years: 11.8%; 3 years: 22.1%; 4 years: 32.4%; 5 years: 39.8%. RFA alone: 1 year: 14.8%; 2 years: 25.0%; 3 years: 35.3%; 4 years: 47.1%; 5 years: 57.4%. RFA + iodine-125: Number of events: 27. RFA alone: Number of events: 39. Comparison between groups: HR 0.508, 95% CI 0.317 to 0.815. RFA + iodine-125: Mean time to recurrence: 93 months. RFA alone: Mean time to recurrence: 66.8 months.</p> <p><b>Response:</b> RFA + iodine-125: Complete ablation after one treatment: 56/68; complete response after two treatments: 12/68. RFA alone: Complete ablation after one treatment: 49/68; complete response after two treatments: 19/68.</p> <p><b>Adverse events:</b> RFA + iodine-125: AEs grade ≥ 3: 15 events (not patient numbers) RFA alone: AEs grade ≥ 3: 11 events (not patient numbers) No procedure-related mortalities and no iodine-125 seed migration from the liver to the heart or other organs.</p>	Low
<b>RFA vs. RFA + chemotherapy</b>					
Gan, 2004 (reported in Chinese) <sup>47</sup> China	38 patients with tumours ≤ 3 cm	RFA alone (n = 18 patients)	RFA + chemotherapy (n = 20 patients)	<p><b>Recurrence:</b> RFA alone: 1 year: 50% (6/12 analysed). RFA + chemotherapy: 1 year: 27% (4/15 analysed). RFA: Number of events: 6 (1 original location recurrence + 5 other location recurrence). RFA + chemotherapy: Number of events: 4 (4 other location recurrence).</p> <p><b>Adverse events:</b> There were no serious AEs in either group.</p>	High

continued

Study name and location	Participant information	Intervention	Comparator	Main results	ROB
<b>RFA + TACE vs. resection</b>					
Liu, 2016 <sup>72</sup> China	200 patients with tumours ≤ 5 cm; subgroup of 135 patients with tumours ≤ 3 cm	Partial hepatectomy (n = 66 patients with tumours ≤ 3 cm)	TACE + RFA (n = 69 patients with tumours ≤ 3 cm)	No results reported for ≤ 3 cm subgroup, other than Kaplan–Meier curve. <b>Adverse events:</b> No 30- or 90-day mortality in either group.	Some concerns
<b>PEI vs. PAI</b>					
Lin, 2005 <sup>53</sup> Taiwan	187 patients with tumours ≤ 3 cm	RFA (n = 62 patients)	PEI (n = 62 patients) PAI (n = 63 patients)	<b>OS:</b> RFA: 1 year: 93%; 2 years: 81%; 3 years: 74%. PEI: 1 year: 88%; 2 years: 66%; 3 years: 51%. PAI: 1 year: 90%; 2 years: 67%; 3 years: 53%. RFA: Number of events (death): 10 (16.1%). PEI: Number of events (death): 17 (27.4%). PAI: Number of events (death): 15 (23.8%). <b>Cancer-free survival:</b> RFA: 1 year: 74%; 2 years: 60%; 3 years: 43%. PEI: 1 year: 70%; 2 years: 41%; 3 years: 21%. PAI: 1 year: 71%; 2 years: 43%; 3 years: 23%. <b>Recurrence:</b> RFA: 1 year: 10%; 2 years: 14%; 3 years: 14%. PEI: 1 year: 16%; 2 years: 34%; 3 years: 34%. PAI: 1 year: 14%; 2 years: 31%; 3 years: 31%. RFA: Number of local recurrence events: 8/60 (13.3%). PEI: Number of local recurrence events: 19/55 (34.5%). PAI: Number of local recurrence events: 17/58 (29.3%). <b>Complete response:</b> RFA: 96.1% (75/78 tumours). PEI: 88.1% (67/76 tumours). PAI: 92.4% (73/79 tumours). <b>Adverse events:</b> RFA: 4.8% (3/62 patients) had serious AEs (2 patients with haemothorax and 1 with gastric bleeding and perforation). PEI: No serious AEs. PAI: No serious AEs. <b>Economic outcomes:</b> RFA: Mean hospitalisation: 4.2 days (range 3–18). PEI: Mean hospitalisation: 1.7 days (range 2–3). PAI: Mean hospitalisation: 2.2 days (range 2–5).	Some concerns

Study name and location	Participant information	Intervention	Comparator	Main results	ROB
<b>PEI vs. resection</b>					
Huang, 2005 <sup>49</sup> Taiwan	82 patients with tumours ≤ 3 cm	PEI (n = 40 patients)	Resection (n = 42 patients)	<p><b>OS:</b> PEI: 1 year: 100%; 2 years: 100%; 3 years: 96.7%; 4 years: 92.1%; 5 years: 46.0%. Resection: 1 year: 97.4%; 2 years: 91.3%; 3 years: 88.1%; 4 years: 88.1%; 5 years: 81.8%.</p> <p><b>PFS:</b> PEI: 1 year: 76.1%; 2 years: 64.5%; 3 years: 49.1%; 4 years: 44.6%; 5 years: 44.6%. Resection: 1 year: 89.5%; 2 years: 71.3%; 3 years: 60.9%; 4 years: 56.2%; 5 years: 48.2%.</p> <p><b>Recurrence:</b> PEI: Number of events: 18. Resection: Number of events: 15.</p> <p><b>Adverse events:</b> PEI: 3 patients had adverse effects (1 decreased blood pressure, 2 wound pain). Resection: No significant complications.</p>	High
<b>PEI vs. RFA + PEI</b>					
Azab, 2011 <sup>45</sup> Egypt	90 patients (with 98 tumours) ≤ 5 cm; subgroup of 48 tumours ≤ 3 cm	PEI + RFA (n = 16 tumours ≤ 3 cm)	RFA alone (n = 16 tumours ≤ 3 cm) PEI alone (n = 16 tumours ≤ 3 cm)	<p><b>Response:</b> PEI + RFA: 15/16 (93.8%) nodules had complete ablation and 1/16 (6.2%) had partial ablation after one session; 16/16 (100%) nodules had complete ablation after two sessions. RFA alone: 12/16 (75%) nodules had complete ablation and 4/16 (25%) had partial ablation after one session; 14/16 (87.5%) nodules had complete ablation and 2/16 (12.5%) had partial ablation after two sessions. PEI alone: 0% nodules had complete ablation after two sessions; 13/16 (81.25%) nodules had complete ablation and 3/16 (18.75%) had partial ablation after all sessions.</p> <p><b>Adverse events:</b> There were no mortalities related to any of the techniques. No other results reported for ≤ 3 cm tumour subgroup.</p>	Some concerns

continued

Study name and location	Participant information	Intervention	Comparator	Main results	ROB
PEI vs. PEI + TACE					
Koda, 2001 <sup>52</sup> Japan	52 patients with tumours < 3 cm	TACE + PEI (n = 26 patients)	PEI alone (n = 26 patients)	<p><b>OS:</b> TACE + PEI: 1 year: 100%; 2 years: 94.4%; 3 years: 80.8%; 5 years: 40.4%. PEI alone: 1 year: 91.3%; 2 years: 81.6%; 3 years: 65.9%; 5 years: 37.7%. TACE + PEI: Number of events: 4. PEI alone: Number of events: 8.</p> <p><b>Recurrence:</b> TACE + PEI local residual disease: 1 year: 3.7%; 2 years: 19.3%; 3 years: 19.3%, 5 years: 19.3%. TACE + PEI new nodular recurrence: 1 year: 8.7%; 2 years: 41.9%; 3 years: 50.2%; 5 years: 50.2%. PEI alone local residual disease: 1 year: 34.2%; 2 years: 39.3%; 3 years: 39.3%, 5 years: 39.3%. PEI alone new nodular recurrence: 1 year: 26.9%; 2 years: 60.1%; 3 years: 80.1%; 5 years: 100%.</p> <p>TACE + PEI local residual disease: Number of events: 4 nodules (of 31). PEI alone local residual disease: Number of events: 11 nodules (of 34). TACE + PEI new nodular recurrence: Number of events: 9. PEI alone new nodular recurrence: Number of events: 17.</p> <p><b>Adverse events:</b> TACE + PEI: Major complications: 2. PEI alone: Major complications: 0. TACE + PEI: After TACE: continuous abdominal pain: 6; severe abdominal pain: 6; fever: 19; high-grade fever: 10; liver dysfunction: 3; leukocytosis: 11; C-reactive protein: 22. After PEI: continuous abdominal pain: 16; severe abdominal pain: 7; fever: 20; high-grade fever: 10; liver dysfunction: 3; leukocytosis: 4; C-reactive protein: 20. PEI alone: continuous abdominal pain: 11; severe abdominal pain: 5; fever: 22; high-grade fever: 5; liver dysfunction: 3; leukocytosis: 5; C-reactive protein: 15.</p>	Low
Mizuki, 2010 <sup>60</sup> Japan	30 patients with tumours ≤ 4 cm (large portion had tumours ≤ 3 cm)	PEI alone (n = 14 patients with average tumour size 2.64 cm)	TACE + PEI (n = 16 patients with average tumour size 2.65 cm)	<p><b>OS:</b> PEI alone: Mean 57.2 months, 95% CI 37.2 to 77.2 months. TACE + PEI: Mean 42.4 months, 95% CI 29.2 to 55.6 months. PEI alone: Number of events (death): 6/14 (44%). TACE + PEI: Number of events (death): 8/13 (61.5%).</p> <p><b>Cancer-free survival:</b> PEI alone: Mean 16.7 months, 95% CI 7.3 to 26.0 months. TACE + PEI: Mean 22.9 months, 95% CI 12.4 to 33.4 months.</p> <p><b>Recurrence:</b> PEI alone: Number of events: 10 (71.4%). TACE + PEI: Number of events: 11 (84.6%).</p> <p><b>Adverse events:</b> In all 30 cases, serious adverse effects or complications were not related to treatment with TACE and/or PEI.</p>	Some concerns

Study name and location	Participant information	Intervention	Comparator	Main results	ROB
<b>PAI vs. PAI + TACE</b>					
Huo, 2003 <sup>70</sup> Taiwan	108 patients with tumours ≤ 5 cm; subgroup of 55 patients with tumours ≤ 3 cm	Sequential TACE and PAI (n = 24 patients with tumours ≤ 3 cm)	PAI alone (n = 31 patients with tumours ≤ 3 cm)	<p><b>OS:</b> Sequential TACE and PAI: 1 year: 100%; 3 years: 73%. PAI alone: 1 year: 100%; 3 years: 49%.</p> <p><b>Cancer-free survival:</b> There were no significant differences in cancer-free survival between the two groups of patients with tumour sizes ≤ 3 cm (p = 0.217).</p> <p><b>Adverse events:</b> Sequential TACE and PAI: Transient fever, abdominal pain and elevation of liver enzymes were present in the majority of patients (in whole sample) after TACE. No serious complications that necessitated intensive care. PAI alone: Side effects relatively mild. Most patients (in whole sample) experienced transient mild to moderate local pain during or after acetic acid injection, which could be controlled with additional analgesics. No serious complications that necessitated intensive care.</p>	High
<b>Percutaneous local ablative therapy vs. resection</b>					
Chen, 2006 <sup>68</sup> China	180 patients with tumours ≤ 5 cm; subgroup of 79 patients with tumours ≤ 3 cm	Percutaneous local ablative therapy (initial RFA followed by RFA/PEI if residual tumour, and TACE if residual tumour remained) (n = 37 patients with tumours ≤ 3 cm)	Partial hepatectomy (n = 42 patients with tumours ≤ 3 cm)	<p>No results reported for ≤ 3 cm subgroup, other than Kaplan–Meier curve.</p> <p>No significant difference in overall and disease-free survival between the two treatment groups in the ≤ 3 cm subgroup.</p>	High

continued

Study name and location	Participant information	Intervention	Comparator	Main results	ROB
MWA + sorafenib vs. resection	Yan, 2016 <sup>56</sup> China	120 patients with tumours ≤ 3 cm Resection (n = 60 patients)	MWA + sorafenib (n = 60 patients)	<p><b>OS:</b> Resection: Mean 51.2 ± 1.5 months. MWA + sorafenib: Mean 64.6 ± 2.4 months. Resection: 1 year: 90.7%; 3 years: 71.5%; 5 years: 56.7%. MWA + sorafenib: 1 year: 91.1%; 3 years: 72.8%; 5 years: 57.5%.</p> <p><b>Tumour-free survival:</b> Resection: 1 year: 87.8%; 3 years: 44.3%; 5 years: 33.2%. MWA + sorafenib: 1 year: 86.2%; 3 years: 48.3%; 5 years: 34.6%.</p> <p><b>Recurrence:</b> Resection: 5 years: 18.3%. MWA + sorafenib: 5 years: 38.3%.</p> <p>Resection: Number of events: 11 (2 local recurrence, 9 distant recurrence). MWA + sorafenib: Number of events: 23 (7 local recurrence, 16 distant recurrence).</p> <p><b>Adverse events:</b> Resection: Pain: 38 (63.3%); fever: 29 (48.3%); abdominal bleeding: 7 (11%); infection: 18 (30%). MWA + sorafenib: Pain: 14 (23.3%); fever: 15 (25%); abdominal bleeding: 2 (3.3%); infection: 1 (1.7%).</p>	High
ICER, incremental cost-effectiveness ratio.					

## Appendix 5 Studies excluded at full paper stage with rationale (non-randomised controlled trial searches)

Study	Reason for exclusion
Abdelaziz, 2014 <sup>64</sup>	Duplicate report
Aikata, 2018 <sup>251</sup>	Not a prospective comparative study
Al-Judaibi, 2018 <sup>252</sup>	Not a prospective comparative study
Anand, 2012 <sup>253</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Barabino, 2018 <sup>254</sup>	Not a prospective comparative study
Bassanello, 2003 <sup>255</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Beyer, 2018 <sup>256</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Bhutiani, 2016 <sup>257</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Borzio, 2007 <sup>258</sup>	No relevant intervention/comparison
Bouda, 2020 <sup>259</sup>	Not a prospective comparative study
Bu, 2015 <sup>260</sup>	Not a prospective comparative study
Bujold, 2011 <sup>261</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Casaccia, 2017 <sup>262</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Chagnon, 2007 <sup>263</sup>	Not a prospective comparative study
Chen, 2007 <sup>264</sup>	No relevant intervention/comparison
Chen, 2019 <sup>265</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Cheung, 2012 <sup>266</sup>	Not a prospective comparative study
Cheung, 2013 <sup>267</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Cho, 2015 <sup>268</sup>	Not a prospective comparative study
Chong, 2020 <sup>269</sup>	Not a prospective comparative study
Cillo, 2014 <sup>270</sup>	Not a prospective comparative study
Costa, 2015 <sup>271</sup>	Not a prospective comparative study
De Geus, 2018 <sup>272</sup>	Not a prospective comparative study
Denecke, 2015 <sup>273</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Di Costanzo, 2015 <sup>127</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Di Sandro, 2019 <sup>274</sup>	Not a prospective comparative study
Ding, 2021 <sup>275</sup>	Not a prospective comparative study
Ei, 2013 <sup>276</sup>	Duplicate report
Eloubeidi, 2000 <sup>277</sup>	Not a prospective comparative study
Ferrari, 2007 <sup>77</sup>	Duplicate report
Freeman, 2021 <sup>278</sup>	Not a prospective comparative study

continued

Study	Reason for exclusion
Gannon, 2009 <sup>279</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Garavoglia, 2013 <sup>280</sup>	No relevant intervention/comparison
Ghweil, 2019 <sup>281</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Guibal, 2013 <sup>282</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Guo, 2005 <sup>140</sup>	Duplicate report
Guo, 2010 <sup>283</sup>	Not a prospective comparative study
Guo, 2013 <sup>141</sup>	Not a prospective comparative study
Hara, 2018 <sup>284</sup>	Not a prospective comparative study
Hara, 2019 <sup>285</sup>	Not a prospective comparative study
Hasegawa, 2013 <sup>286</sup>	Not a prospective comparative study
He, 2018 <sup>287</sup>	Not a prospective comparative study
Helmberger, 2007 <sup>288</sup>	No relevant intervention/comparison
Hiraoka, 2008 <sup>289</sup>	Not a prospective comparative study
Hiraoka, 2017 <sup>290</sup>	Not a prospective comparative study
Ho, 2012 <sup>291</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Hong, 2005 <sup>145</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Hsiao, 2020 <sup>146</sup>	Not a prospective comparative study
Hsu, 2013 <sup>292</sup>	Not a prospective comparative study
Huang, 2012 <sup>293</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Hung, 2011 <sup>150</sup>	Not a prospective comparative study
Iida, 2014 <sup>152</sup>	Not a prospective comparative study
Ikeda, 2005 <sup>294</sup>	Not a prospective comparative study
Ikeda, 2007 <sup>295</sup>	Not a prospective comparative study
Imai, 2012 <sup>296</sup>	Duplicate report
Imai, 2013	Not a prospective comparative study
Ismailova, 2017	Not early HCC patients ( $\leq 3$ cm tumour)
Ito, 2015 <sup>297</sup>	Duplicate report
Ito, 2016 <sup>298</sup>	Not a prospective comparative study
Jiang, 2015	Not a prospective comparative study
Jianyong, 2017 <sup>299</sup>	Not a prospective comparative study
Juloori, 2020 <sup>300</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Kang, 2014 <sup>301</sup>	Not a prospective comparative study
Kanwal, 2012 <sup>302</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Katsoulakis, 2014 <sup>303</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Kawamura, 2019 <sup>304</sup>	Not a prospective comparative study
Kawaoka, 2019 <sup>305</sup>	Not a prospective comparative study
Kennedy, 2004 <sup>306</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Kim, 2014a <sup>307</sup>	Not early HCC patients ( $\leq 3$ cm tumour)

Study	Reason for exclusion
Kim, 2014b <sup>308</sup>	Not a prospective comparative study
Kim, 2016 <sup>309</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Kim, 2018 <sup>310</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Kim, 2019 <sup>311</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Kim, 2021 <sup>162</sup>	Duplicate report
Kimura, 2018a <sup>312</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Kimura, 2018b <sup>313</sup>	Duplicate report
Ko, 2020 <sup>314</sup>	Not a prospective comparative study
Komatsu, 2011 <sup>315</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Komatsu, 2019 <sup>316</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Kooby, 2010 <sup>317</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Kuang, 2011 <sup>318</sup>	Not a prospective comparative study
Kudithipudi, 2017 <sup>319</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Kuo, 2010 <sup>320</sup>	Not a prospective comparative study
Kuo, 2021 <sup>321</sup>	Not a prospective comparative study
Kuromatsu, 2009 <sup>322</sup>	Not a prospective comparative study
Kwon, 2012 <sup>323</sup>	No relevant intervention/comparison
Lai, 2016 <sup>169</sup>	Not a prospective comparative study
Lapinski, 2021 <sup>324</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Lee, 2012 <sup>325</sup>	Not a prospective comparative study
Lee, 2014 <sup>326</sup>	Not a prospective comparative study
Lee, 2018 <sup>327</sup>	Not a prospective comparative study
Lee, 2019 <sup>328</sup>	Duplicate report
Lee, 2021 <sup>329</sup>	Not a prospective comparative study
Lei, 2014 <sup>330</sup>	Not a prospective comparative study
Li, 2015 <sup>331</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Li, 2016 <sup>332</sup>	Not a prospective comparative study
Li, 2019 <sup>333</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Lin, 2007 <sup>179</sup>	Not a prospective comparative study
Liu, 2016 <sup>334</sup>	Not a prospective comparative study
Liu, 2017 <sup>335</sup>	Not a prospective comparative study
Liu, 2018a <sup>336</sup>	Not a prospective comparative study
Liu, 2018b <sup>337</sup>	Not a prospective comparative study
Liu, 2019 <sup>338</sup>	Not a prospective comparative study
Loo, 2020 <sup>339</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Lu, 2010 <sup>340</sup>	Not a prospective comparative study
Maezawa, 2005 <sup>341</sup>	Not early HCC patients ( $\leq 3$ cm tumour)

continued

Study	Reason for exclusion
Mariani, 2017 <sup>342</sup>	Not a prospective comparative study
Martin, 2016 <sup>343</sup>	Not a prospective comparative study
Merle, 2005 <sup>344</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Ming, 2012 <sup>345</sup>	Not a prospective comparative study
Mizumoto, 2011 <sup>346</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Mo, 2003 <sup>347</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Mohnike, 2019 <sup>348</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Molinari, 2009 <sup>349</sup>	Duplicate report
Molinari, 2014 <sup>350</sup>	Not a prospective comparative study
Mornex, 2005 <sup>351</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Mornex, 2007 <sup>197</sup>	Not a prospective comparative study
Murakami, 2007 <sup>198</sup>	Not a prospective comparative study
Nahon, 2021 <sup>352</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Nanashima, 2004 <sup>353</sup>	Not a prospective comparative study
Nanashima, 2010 <sup>354</sup>	Not a prospective comparative study
Nathan, 2013 <sup>355</sup>	Not a prospective comparative study
Ogiso, 2021 <sup>356</sup>	Not a prospective comparative study
Oh, 2019 <sup>357</sup>	Not a prospective comparative study
Oh, 2020 <sup>358</sup>	Not a prospective comparative study
Ohmoto, 2006 <sup>359</sup>	Not a prospective comparative study
Paik, 2016 <sup>205</sup>	Not a prospective comparative study
Pan, 2020 <sup>360</sup>	Not a prospective comparative study
Park, 2007 <sup>361</sup>	Not a prospective comparative study
Park, 2018 <sup>362</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Park, 2020 <sup>363</sup>	Duplicate report
Peng, 2011 <sup>364</sup>	Not a prospective comparative study
Peng, 2012 <sup>365</sup>	Duplicate report
Peng, 2013 <sup>366</sup>	Not a prospective comparative study
Peng, 2014 <sup>367</sup>	Not a prospective comparative study
Pompili, 2013 <sup>210</sup>	Not a prospective comparative study
Praktiknjo, 2018 <sup>368</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Pryor, 2019 <sup>369</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Rong, 2020 <sup>370</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Ruzzenente, 2012 <sup>371</sup>	Not a prospective comparative study
Ryu, 2018 <sup>372</sup>	Not a prospective comparative study
Ryu, 2019 <sup>373</sup>	Not a prospective comparative study
Sako, 2003 <sup>374</sup>	Not a prospective comparative study
Santambrogio, 2009a <sup>214</sup>	Not early HCC patients ( $\leq 3$ cm tumour)

Study	Reason for exclusion
Santambrogio, 2009b <sup>375</sup>	Not a prospective comparative study
Santambrogio, 2017 <sup>376</sup>	Not a prospective comparative study
Santambrogio, 2018 <sup>377</sup>	Not a prospective comparative study
Santambrogio, 2021 <sup>378</sup>	Not a prospective comparative study
Schaible, 2020 <sup>379</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Sheng, 2015 <sup>380</sup>	Not a prospective comparative study
Shi, 2014 <sup>217</sup>	Not a prospective comparative study
Shibata, 2001 <sup>381</sup>	Duplicate report
Shiozawa, 2015 <sup>382</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Simo, 2011 <sup>383</sup>	Not a prospective comparative study
Song, 2019 <sup>384</sup>	Not a prospective comparative study
Spangenberg, 2008 <sup>385</sup>	Not a prospective comparative study
Stuart, 2018 <sup>386</sup>	Not a prospective comparative study
Su, 2020 <sup>387</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Suh, 2013 <sup>388</sup>	Not a prospective comparative study
Sun, 2020 <sup>219</sup>	Not a prospective comparative study
Sutter, 2018 <sup>389</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Takamatsu, 2014 <sup>390</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Takami, 2009 <sup>391</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Takami, 2010 <sup>392</sup>	Not a prospective comparative study
Takayama, 2010 <sup>393</sup>	Not a prospective comparative study
Takayasu, 2018a <sup>394</sup>	Not a prospective comparative study
Takayasu, 2018b <sup>394</sup>	Duplicate report
Takeda, 2008 <sup>395</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Takeda, 2016 <sup>396</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Tanaka, 2015 <sup>397</sup>	Not a prospective comparative study
Tanguturi, 2015 <sup>398</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Tashiro, 2011 <sup>220</sup>	Not a prospective comparative study
Tatineni, 2019 <sup>399</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Teramoto, 2005 <sup>400</sup>	Not a prospective comparative study
Toro, 2012 <sup>401</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Toyoda, 2008 <sup>221</sup>	Not a prospective comparative study
Trotschel, 2016 <sup>402</sup>	Not a prospective comparative study
Ueno, 2020 <sup>403</sup>	Not a prospective comparative study
Utsunomiya, 2014 <sup>404</sup>	Not a prospective comparative study
Vietti Violi, 2017a <sup>405</sup>	Duplicate report
Vietti Violi, 2017b <sup>406</sup>	Duplicate report

continued

Study	Reason for exclusion
Vitale, 2012 <sup>407</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Vitali, 2016 <sup>408</sup>	Not a prospective comparative study
Vivarelli, 2004 <sup>223</sup>	Not a prospective comparative study
Wang, 2007 <sup>409</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Wang, 2012 <sup>410</sup>	Not a prospective comparative study
Wang, 2015a <sup>411</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Wang, 2015b <sup>412</sup>	Not a prospective comparative study
Wang, 2019 <sup>413</sup>	Not a prospective comparative study
Wang, 2020 <sup>414</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Wei, 2020 <sup>415</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Wigg, 2017 <sup>416</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Wiggermann, 2012 <sup>417</sup>	Not a prospective comparative study
Wong, 2021 <sup>418</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Wu, 2020 <sup>419</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Xie, 2019 <sup>420</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Xu, 2009 <sup>421</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Xu, 2014 <sup>422</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Xu, 2015 <sup>423</sup>	Not a prospective comparative study
Yamao, 2018 <sup>424</sup>	Not a prospective comparative study
Yamashita, 2017 <sup>425</sup>	Not a prospective comparative study
Yamashita, 2019 <sup>426</sup>	Not a prospective comparative study
Yamazaki, 2009 <sup>427</sup>	Not a prospective comparative study
Yang, 2010 <sup>428</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Ye, 2008 <sup>429</sup>	No relevant intervention/comparison
Yi, 2014 <sup>430</sup>	No relevant intervention/comparison
Yohji, 2012 <sup>431</sup>	Not a prospective comparative study
Yoon, 2018 <sup>432</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Yu, 2014 <sup>433</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Yun, 2009 <sup>434</sup>	Duplicate report
Yun, 2011 <sup>239</sup>	Not a prospective comparative study
Zhang, 2008 <sup>435</sup>	No relevant intervention/comparison
Zhang, 2013 <sup>241</sup>	Not a prospective comparative study
Zhang, 2016 <sup>243</sup>	Not a prospective comparative study
Zheng, 2020 <sup>436</sup>	Not a prospective comparative study
Zhou, 2014 <sup>247</sup>	Not a prospective comparative study
Zhu, 2007 <sup>249</sup>	Not early HCC patients ( $\leq 3$ cm tumour)
Zhu, 2021 <sup>80</sup>	Duplicate report

## **Appendix 6** Characteristics and results of non-randomised studies included in the review

Study name and location	Participant information	Intervention	Comparator	Main results	ROB
<b>RFA vs. MWA</b>					
Barabino, 2016 <sup>83</sup> (conference abstract) Italy	154 patients with HCC unsuitable for percutaneous or hepatic resection	Laparoscopic RFA [n = 94 patients with average tumour size 1.92 (±0.5) cm]	Laparoscopic MWA [n = 60 patients with average tumour size 2.15 (±0.53) cm]	<p><b>OS:</b> Laparoscopic RFA: 5 years: 50%. Laparoscopic MWA: 5 years: 37%.</p> <p><b>Disease-free survival:</b> Laparoscopic RFA: 5 years: 19%. Laparoscopic MWA: 5 years: 12%. Laparoscopic RFA: Local tumour progression rate: 21.2%. Laparoscopic MWA: Local tumour progression rate: 8.3%.</p> <p><b>Response:</b> Laparoscopic RFA: Complete ablation: 95%. Laparoscopic MWA: Complete ablation: 95%.</p> <p><b>Adverse events:</b> Laparoscopic RFA: 1% major complications. Laparoscopic MWA: 2% major complications. No deaths related to the procedure in either group.</p>	High
Qian, 2012 <sup>89</sup> China	42 patients with < 3 cm tumours	MWA (n = 22 patients)	RFA (n = 20 patients)	<p><b>PFS:</b> MWA: 4/22 (18.2%) had local tumour progression, 1/22 (4.5%) had a new intrahepatic tumour. RFA: 3/20 (15%) had local tumour progression, 4/20 (20%) had a new intrahepatic tumour.</p> <p><b>Response:</b> MWA: Complete ablation: 21/22 (95.5%) after the first ablation, 100% after the second ablation. RFA: Complete ablation: 19/20 (95%) after the first ablation, 100% after the second ablation.</p> <p><b>Adverse events:</b> No skin burns, tumour seeding or treatment-related death in either group.</p>	Low
<b>RFA vs. resection</b>					
Choi, 2004 <sup>84</sup> (conference abstract) Korea	164 patients with ≤ 3 cm tumours	RFA (n = 62 patients)	Hepatic resection (n = 102 patients)	<p><b>OS:</b> RFA: 1 year: 100%; 3 years: 73.9%. Resection: 1 year: 97.1%; 3 years: 83.0%.</p> <p><b>Recurrence-free survival:</b> RFA: 1 year: 74.1%; 3 years: 40.2%. Resection: 1 year: 75.9%; 3 years: 54.7%. RFA: Local recurrence rate: 11.3%; Remote recurrence rate: 53.7%. Resection: Local recurrence rate: 2.0%; Remote recurrence rate: 45.3%.</p>	High

Study name and location	Participant information	Intervention	Comparator	Main results	ROB
Du, 2012 <sup>22</sup> China	116 patients with tumours ≤ 5 cm; subgroup of 60 patients with tumours ≤ 3 cm	RFA (n = 31 patients with tumours ≤ 3 cm)	Surgical resection (n = 29 patients with tumours ≤ 3 cm)	<b>Relapse:</b> RFA: 1 year: 12.9%. Resection: 1 year: 13.8%. <b>RFA:</b> Number of events: 4. <b>Resection:</b> Number of events: 4. No other results reported for ≤ 3cm tumour subgroup.	High
Elgendy, 2014 <sup>85</sup> (conference abstract) Egypt	51 patients with < 3 cm tumours in locations not amenable for percutaneous route	Intraoperative RFA (n = 24 patients)	Surgical resection (n = 27 patients)	<b>OS:</b> Intraoperative RFA: 1 year: 92%; 3 years: 74%. Resection: 1 year: 93%; 3 years: 81%. <b>Recurrence:</b> No tumours showed local recurrence after median 37 months follow-up in either group. <b>Response:</b> Intraoperative RFA: Complete ablation: 100%. Resection: Complete resection: 100%. <b>Adverse events:</b> Complication rate was comparable between treatment groups.	High
Elgendy, 2015 <sup>71</sup> (conference abstract) Egypt	92 patients with < 2 cm tumours in locations not amenable for percutaneous route	Intraoperative RFA (n = 44 patients)	Surgical resection (n = 48 patients)	<b>OS:</b> Intraoperative RFA: 1 year: 91%; 3 years: 76%. Resection: 1 year: 92%; 3 years: 83%. <b>Recurrence:</b> No tumours showed local recurrence after median 46 months follow-up in either group. <b>Response:</b> Intraoperative RFA: Complete ablation: 100%. Resection: Complete resection: 100%. <b>Adverse events:</b> Complication rate was comparable between treatment groups.	High

continued

Study name and location	Participant information	Intervention	Comparator	Main results	ROB
Harada, 2016 <sup>6</sup> Japan	121 patients with < 5 cm tumours and portal hypertension (large proportion had tumours ≤ 3 cm)	RFA (n = 40 patients)	Surgical resection [n = 81 patients, a few had tumours > 3 cm; mean 2.1 cm (range 0.7–5 cm)]	<p><b>OS:</b> RFA: 1 year: 97.2%; 3 years: 84.1%; 5 years: 50.6%. Resection: 1 year: 93.4%; 3 years: 84.5%; 5 years: 37.1%. HR (RFA vs. resection) 0.86, 95% CI 0.58 to 1.74.</p> <p><b>Recurrence-free survival:</b> RFA: 1 year: 36.1%; 3 years: 22.2%; 5 years: 4.8%. Resection: 1 year: 74.1%; 3 years: 48.8%; 5 years: 42.9%. RFA: Number of events: 34/40 (85%); 27 intrahepatic distant recurrences and 7 local recurrences. Resection: Number of events: 34/81 (42%); 33 intrahepatic distant recurrences and one local recurrence. HR (RFA vs. resection) 2.74, 95% CI 1.70 to 4.43.</p> <p><b>Adverse events:</b> RFA: 5 (12.5%) Resection: 38 (46.9%) RFA: Clavien–Dindo ≥ grade III AEs: 1 (2.5%). Resection: Clavien–Dindo ≥ grade III AEs: 13 (16.1%). One in-hospital death occurred in the resection group, secondary to sepsis.</p> <p><b>Economic outcomes:</b> RFA: Median postoperative hospital stay: 8 days (range 2–29). Resection: Median postoperative hospital stay: 15 days (range 7–80).</p>	High
Huang, 2014 <sup>87</sup> China	346 patients with ≤ 3 cm tumours	RFA (n = 121 patients)	Surgical resection (n = 225 patients)	<p><b>Health-related quality of life:</b> FACT-Hep: Patients treated with RFA had significantly better HRQL total scores after 3, 6, 12, 24 and 36 months than those who had resection (further details reported in supplementary tables/figures).</p> <p><b>Adverse events:</b> RFA: 6 Clavien–Dindo grade III (2 pleural effusion requiring tapping, 2 pneumothorax or haemothorax, 1 liver abscess, 1 intra-abdominal haemorrhage). No grade I, II, IV or V AEs. Resection: 10 Clavien–Dindo grade I, 3 Clavien–Dindo grade II, 59 Clavien–Dindo grade III (9 delayed gastric emptying, 14 pleural effusion requiring tapping, 10 biliary leakage, 5 intra-abdominal abscess, 15 moderate/severe ascites, 6 intra-abdominal haemorrhage), 3 grade IV (2 liver failure, 1 renal insufficiency), 0 grade V AEs. No hospital death in either group.</p> <p><b>Economic outcomes:</b> RFA: Average duration of procedure: 44 minutes. Resection: Average duration of procedure: 166.5 minutes. RFA: Average length of hospital stay: 7.1 days. Resection: Average length of hospital stay: 14.5 days. Blood transfusion rates were higher in the resection group. No other results reported, other than Kaplan–Meier curves.</p>	Unclear

Study name and location	Participant information	Intervention	Comparator	Main results	ROB
Peng, 2010 <sup>88</sup> (conference abstract) China	195 patients with ≤ 3 cm tumours	RFA (n = 79) as first choice, otherwise surgical resection (n = 24) (total n = 103 patients)	Surgical resection (n = 75) as first choice, otherwise RFA (n = 17) (total n = 92 patients)	<b>OS:</b> RFA: 1 year: 95.9%; 3 years: 75.8%; 4 years: 70.7%. Resection: 1 year: 90.1%; 3 years: 63.7%; 4 years: 55.5%. <b>Adverse events:</b> RFA: No cases of treatment-related mortality. Resection: Two cases of treatment-related mortality.	High
Tateishi, 2020 (conference abstract) <sup>90</sup> Japan	740 patients with ≤ 3 cm tumours	RFA (n = 369 patients)	Surgery (n = 371 patients)	<b>Recurrence:</b> RFA: 3 years: 61.7%. Surgery: 3 years: 66.0%. RFA: Number of events: 218. Surgery: Number of events: 192. Adjusted HR using inverted probability of treatment weighting 0.89, 95% CI 0.72 to 1.1.	High
<b>MWA vs. resection</b>					
Horigome, 2000 <sup>86</sup> Japan	105 patients with ≤ 3 cm tumours	Hepatic resection (n = 43 patients)	MWA (n = 29 patients) PEI (n = 33 patients)	<b>Recurrence:</b> Resection: Number of events: 31 (72%); 13 had intrahepatic metastases, 18 had multicentric occurrence. MWA: Number of events: 11 (38%); 7 had intrahepatic metastases, 4 had multicentric occurrence. PEI: Number of events: 22 (67%); 16 had intrahepatic metastases, 6 had multicentric occurrence.	High
<b>HIFU vs. RFA</b>					
Cheung, 2013 <sup>83</sup> China	106 patients (with 119 tumours) with < 3 cm tumours (primary or first recurrence)	HIFU (n = 47 patients, 52 tumours)	RFA (n = 59 patients, 67 tumours)	<b>OS:</b> HIFU: 1 year: 97.4%, 3 years: 81.2%. RFA: 1 year: 94.6%, 3 years: 79.8%. <b>Disease-free survival:</b> HIFU: 1 year: 63.6%, 3 years: 25.9%. RFA: 1 year: 62.4%, 3 years: 34.1%. <b>Response:</b> HIFU: Complete response: 87.2%. RFA: Complete response: 94.9%. <b>Adverse events:</b> HIFU: 10 (21.3%); 2 pneumothorax, 1 third degree skin burn, 7 relatively minor complications. RFA: 5 (8.5%); 2 pleural effusion requiring tapping, 1 liver abscess, 1 oesophageal variceal bleeding, 1 mild wound infection. HIFU: Clavien–Dindo ≥ grade III AEs: 3. RFA: Clavien–Dindo ≥ grade III AEs: 4. <b>Economic outcomes:</b> HIFU: Median hospital stay: 4 days (range 2–18). RFA: Median hospital stay: 6 days (range 1–31).	High

continued

Study name and location	Participant information	Intervention	Comparator	Main results	ROB
<b>Cryoablation vs. RFA or MWA</b>					
Ei, 2015 <sup>95</sup> Japan	119 patients with < 5 cm tumours (large proportion had tumours ≤ 3 cm)	Cryoablation (n = 55 patients with median tumour size 2.5 cm; maximum 4 cm)	RFA (n = 27) or microwave (n = 37) (Total n = 64 patients with median tumour size 1.9 cm; maximum 4.5 cm)	<p><b>OS:</b> Cryoablation: 2 years: 88% in ≤ 2 cm subgroup; 86% in &gt; 2 cm subgroup. RFA/MWA: 2 years: 95% in ≤ 2 cm subgroup; 85% in &gt; 2 cm subgroup.</p> <p><b>Local recurrence-free survival:</b> Cryoablation: 2 years: 80%. RFA/MWA: 2 years: 68%.</p> <p><b>Local recurrence:</b> Cryoablation: 2 years: 19% in ≤ 2 cm subgroup; 21% in &gt; 2 cm subgroup. RFA/MWA: 2 years: 23% in ≤ 2 cm subgroup; 56% in &gt; 2 cm subgroup.</p> <p><b>Initial recurrence at other sites of the liver:</b> Cryoablation: 2 years: 38%. RFA/MWA: 2 years: 34%.</p> <p><b>Distant metastases (bone/lung):</b> Cryoablation: 2 years: 2 patients.</p> <p><b>Adverse events:</b> Cryoablation: 6; 3 Clavien–Dindo grade III (2 pleural effusion, 1 intraabdominal bleeding), 3 grade I/II. RFA/MWA: 7; 3 Clavien–Dindo grade III (haemothorax, wound infection, pneumonia), 4 grade I/II. No in-hospital mortality in either group.</p> <p><b>Economic outcomes:</b> Cryoablation: Median operative time: 180 minutes. RFA/MWA: Median operative time: 132 minutes. Length of hospital stay: median 8 days (IQR 6–11 days) in both groups.</p>	High
<b>IRE vs. RFA</b>					
Sugimoto, 2019 <sup>94</sup> Japan	21 patients (with 24 tumours)	IRE [n = 10 patients (13 tumours) with median tumour size 2.03 (SD 0.44) cm]	RFA [n = 11 patients (11 tumours) with median tumour size 1.73 (SD 0.67) cm]	<p><b>Recurrence:</b> IRE: 1 patient had local tumour progression at 6 months. RFA: 0 patients had local tumour progression at 6 months. No other results reported, other than systemic immune responses.</p>	High
IQR, interquartile range.					



EME  
HSDR  
**HTA**  
PGfAR  
PHR

Part of the NIHR Journals Library  
[www.journalslibrary.nihr.ac.uk](http://www.journalslibrary.nihr.ac.uk)

*This report presents independent research funded by the National Institute for Health and Care Research (NIHR).  
The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the  
Department of Health and Social Care*

***Published by the NIHR Journals Library***



OPEN ACCESS

# Management of patients presenting to the emergency department with sudden onset severe headache: systematic review of diagnostic accuracy studies

Matthew Walton ,<sup>1</sup> Robert Hodgson,<sup>1</sup> Alison Eastwood,<sup>1</sup> Melissa Harden,<sup>1</sup> James Storey,<sup>2</sup> Taj Hassan,<sup>3</sup> Marc Stuart Randall ,<sup>4</sup> Abu Hassan,<sup>3</sup> John Williams,<sup>5</sup> Ros Wade<sup>1</sup>

Handling editor Richard Body

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/emmermed-2021-211900>).

<sup>1</sup>Centre for Reviews and Dissemination, University of York, York, UK

<sup>2</sup>Department of Acute Internal Medicine, Leeds Teaching Hospitals NHS Trust, Leeds, UK

<sup>3</sup>Department of Emergency Medicine, Leeds Teaching Hospitals NHS Trust, Leeds, UK

<sup>4</sup>Department of Adult Neurology, Leeds Teaching Hospitals NHS Trust, Leeds, UK

<sup>5</sup>Patient representative, UK, UK

## Correspondence to

Mr Matthew Walton, University of York, York, UK; [matthew.walton@york.ac.uk](mailto:matthew.walton@york.ac.uk)

Received 30 July 2021

Accepted 14 March 2022



© Author(s) (or their employer(s)) 2022. Re-use permitted under CC BY. Published by BMJ.

**To cite:** Walton M, Hodgson R, Eastwood A, et al. *Emerg Med J* Epub ahead of print: [please include Day Month Year]. doi:10.1136/emmermed-2021-211900

## ABSTRACT

**Objective** Advances in imaging technologies have precipitated uncertainty and inconsistency in the management of neurologically intact patients presenting to the Emergency Department (ED) with non-traumatic sudden onset severe headache with a clinical suspicion of subarachnoid haemorrhage (SAH). The objective of this systematic review was to evaluate diagnostic strategies in these patients.

**Methods** Studies assessing any decision rule or diagnostic test for evaluating neurologically intact adults with a severe headache, reaching maximum intensity within 1 hour, were eligible. Eighteen databases (including MEDLINE and Embase) were searched. Quality was assessed using QUADAS-2. Where appropriate, hierarchical bivariate meta-analysis was used to synthesise diagnostic accuracy results.

**Results** Thirty-seven studies were included. Eight studies assessing the Ottawa SAH clinical decision rule were pooled; sensitivity 99.5% (95% CI 90.8 to 100), specificity 24% (95% CI 15.5 to 34.4). Four studies assessing CT within 6 hours of headache onset were pooled; sensitivity 98.7% (95% CI 96.5 to 100), specificity 100% (95% CI 99.7 to 100). The sensitivity of CT beyond 6 hours was considerably lower ( $\leq 90\%$ ; 2 studies). Three studies assessing lumbar puncture (LP; spectrophotometric analysis) following negative CT were pooled; sensitivity 100% (95% CI 100 to 100), specificity 95% (95% CI 86.0 to 98.5).

**Conclusion** The Ottawa SAH Rule rules out further investigation in only a small proportion of patients. CT undertaken within 6 hours (with expertise of a neuroradiologist or radiologist who routinely interprets brain images) is highly accurate and likely to be sufficient to rule out SAH; CT beyond 6 hours is much less sensitive. The CT-LP pathway is highly sensitive for detecting SAH and some alternative diagnoses, although LP results in some false positive results.

## INTRODUCTION

Non-traumatic acute headache accounts for around 2% of adult Emergency Department (ED) attendances.<sup>1</sup> Sudden onset severe headaches may be caused by a primary headache disorder or may be secondary to a more serious underlying pathology, such as subarachnoid haemorrhage (SAH). Diagnosis of SAH is particularly challenging in alert, neurologically intact patients presenting with acute

## Key messages

### What is already known on this subject

- Guidelines typically recommend non-contrast CT head followed by lumbar puncture in patients who present with headache symptoms suspicious for subarachnoid haemorrhage.
- More recently, studies have questioned the need for routine lumbar puncture after a normal CT head.
- Additionally, a decision rule to direct imaging has been widely studied.

### What this study adds

- In this systematic review and meta-analysis, we found that the Ottawa subarachnoid haemorrhage clinical decision rule has low specificity, and could result in significant additional unnecessary testing.
- CT head within 6 hours of headache onset, with images assessed by a neuroradiologist or radiologist who routinely interprets brain images, is highly accurate; around 658 CT-negative patients would have to undergo further investigation to identify a single case of subarachnoid haemorrhage.
- CT head undertaken beyond 6 hours is much less sensitive, therefore additional testing is more likely to be beneficial.
- In healthcare systems and settings in which neuroradiology expertise is unavailable, caution should be exercised when translating the diagnostic accuracy of CT head in the literature to clinical decision making.

### How this study might affect research, practice or policy

- CT head within 6 hours of headache onset and with access to neuroradiology expertise is likely to be sufficient to rule out subarachnoid haemorrhage.
- The diagnostic accuracy of CT head may be contingent on time since symptom onset, which must be accounted for in practice, and investigated in future research.
- Risk tolerance of the patient and physician for the potential consequences of investigation and missed diagnoses will continue to inform practice.

severe headache. Clinical features separating these patients from higher volume complaints with a similar presentation (eg, migraine) are often unreliable indicators of who requires further investigation.<sup>2</sup>

Advances in imaging technologies have precipitated uncertainty and inconsistency in the optimal management of neurologically intact patients presenting to the ED with non-traumatic sudden onset severe headache.<sup>3,4</sup> Given increasing evidence on the potentially low therapeutic value of lumbar puncture (LP) following CT of the head, and its associated adverse effects,<sup>3,5-7</sup> updated evidence-based guidance is needed. We therefore undertook a systematic review of evidence on diagnostic strategies for neurologically intact adult patients presenting to hospital with non-traumatic sudden onset severe headache, reaching maximum intensity within 1 hour.

## METHODS

The review protocol is registered on PROSPERO (CRD42020173265). This paper conforms to the recommendations of the Preferred Reporting Items for a Systematic Review and Meta-Analysis of Diagnostic Test Accuracy Studies statement.<sup>8</sup>

### Search strategy and selection criteria

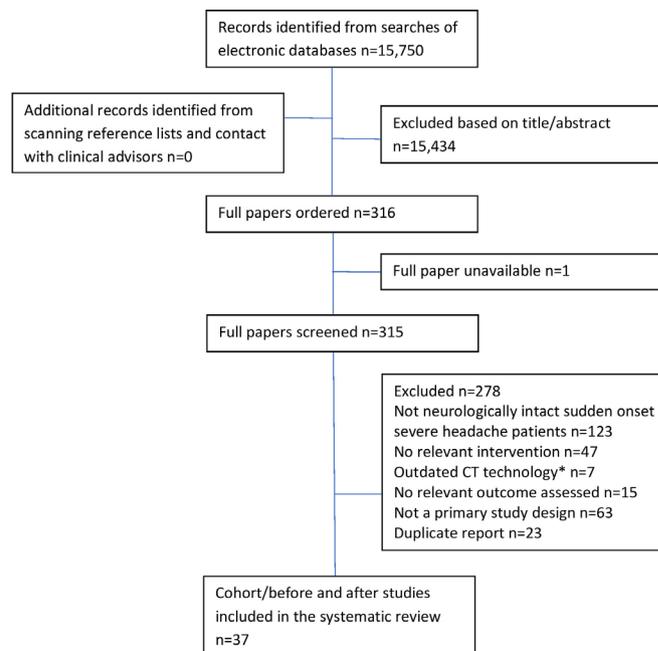
Eighteen databases (including MEDLINE and Embase) were systematically searched in February 2020. Further details of the search strategy are presented in online supplemental file 1. To meet inclusion criteria, studies had to assess any care pathway for ruling out SAH (including clinical decision rules and specific diagnostic tests, such as CT or LP) in neurologically intact adult patients presenting to hospital with a sudden onset severe headache (reaching maximum intensity within 1 hour), with a clinical suspicion of SAH. Studies of patients who had suffered a head injury (ie, traumatic headache) were excluded. Any primary study design (other than single case study) was eligible for inclusion. Outcomes of interest included diagnostic accuracy, quality of life and adverse events. Two researchers (MW and RW) independently screened the titles and abstracts of all retrieved records and subsequently all full text publications for inclusion. Disagreements at each stage of the study selection process were resolved through discussion. Authors of potentially relevant conference abstracts were contacted for additional information. Relevant foreign language studies were translated and included in the review.

### Data extraction and quality assessment

Data were extracted on study methods, patient, intervention and reference standard characteristics, outcome measures, adverse events and results (presented in online supplemental file 2). Data extraction and quality assessment were undertaken by one researcher and independently checked by a second. The majority of studies were assessed for quality using the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) tool.<sup>9</sup> The QUADAS-2 tool was not appropriate for studies where a reference standard test was not used, therefore, a quality assessment tool was developed by RW specifically for the review, piloted and refined before use (see online supplemental file 3 for details).

### Data analysis

Where sufficient information was reported, diagnostic accuracy data were extracted into 2×2 tables to calculate sensitivity, specificity, false positive and false negative rates. Where equivalent diagnostic strategies or tools were used in three or more



**Figure 1** Flow diagram of the study selection process. \*Any study which recruited patients before the year 2000 was considered to have used outdated CT technology.

studies, the hierarchical bivariate model described by Reitsma *et al*<sup>10</sup> was fitted, along with an extension described by Simmonds and Higgins<sup>11</sup> to meta-analyse sensitivity and specificity while accounting for correlation between the two, and within-person correlation between test results. Meta-analyses used standard random-effects DerSimonian-Laird methods. Subgroups were analysed separately to account for underlying differences in diagnostic strategies. The diagnostic accuracy of CT conducted <6 hours from headache onset was analysed separately, as CT accuracy is known to drop rapidly outside of this time frame.<sup>12</sup> The accuracy of different methods of cerebrospinal fluid (CSF) analysis was also assessed. Where results could not be pooled, they were synthesised narratively along with reported adverse event data.

### Public and patient involvement

A patient collaborator with experience of presenting to an ED with a sudden onset severe headache was involved throughout the project. Three additional patients were recruited to an advisory group. The patients provided input during protocol development and interpretation of review findings.

## RESULTS

The search strategy identified 15 750 records; 37 cohort/before and after studies were eligible for inclusion (figure 1 and table 1). More detailed study characteristics and results are presented in online supplemental file 2.

Twelve studies had a low risk of bias for all domains, the other 25 were at risk of bias. Twenty-eight studies were assessed using the QUADAS-2 tool; results are summarised in figure 2.<sup>9</sup> Nine studies did not use a reference standard test, therefore, QUADAS-2 was inappropriate; a quality assessment tool developed specifically for the review was used instead. Quality assessment results are presented in the online supplemental file 3.

**Table 1** Studies included in the systematic review

Intervention	Study	Location	N	Study design
Clinical decision rules (Canadian clinical decision rules 1, 2, 3; Ottawa SAH Rule)	Bellolio <i>et al</i> <sup>13</sup>	USA	454	Retrospective cohort
	Cheung <i>et al</i> <sup>14</sup>	Hong Kong	500	Retrospective cohort
	Chu <i>et al</i> <sup>15</sup>	Australia	137	Retrospective cohort (substudy of a prospective cohort)
	Kelly <i>et al</i> <sup>16</sup>	Australia	59	Retrospective cohort
	MacDonald <i>et al</i> <sup>17</sup>	UK	280	Retrospective cohort
	Matloob <i>et al</i> <sup>18</sup>	UK	112	Retrospective cohort
	Pathan <i>et al</i> <sup>19</sup>	UK	145	Retrospective cohort
	Perry <i>et al</i> <sup>20</sup>	Canada	1999	Prospective cohort
	Perry <i>et al</i> <sup>21</sup>	Canada	2131	Prospective cohort
	Perry <i>et al</i> <sup>22</sup>	Canada	1153; overlap with Perry <i>et al</i> <sup>23</sup>	Prospective cohort
	Perry <i>et al</i> <sup>23</sup>	Canada	3672	Prospective before/after
Wu <i>et al</i> <sup>24</sup>	Taiwan	913	Retrospective cohort	
Yiangou <i>et al</i> <sup>25</sup>	UK	162	Retrospective cohort	
CT-LP pathway	Blok <i>et al</i> <sup>26</sup>	The Netherlands	760	Retrospective cohort
	Cooper <i>et al</i> <sup>7</sup>	UK	517	Retrospective cohort
	Dutto <i>et al</i> <sup>27</sup>	Italy	70	Before/After
	Perry <i>et al</i> <sup>28</sup>	Canada	891	Retrospective cohort
	Perry <i>et al</i> <sup>29</sup>	Canada	592	Prospective cohort
	Valle Alonso <i>et al</i> <sup>30</sup>	Spain	74	Retrospective cohort
CT	Austin <i>et al</i> <sup>31</sup>	UK	250	Retrospective cohort
	Backes <i>et al</i> <sup>32</sup>	The Netherlands	250	Retrospective cohort
	Blok <i>et al</i> <sup>26</sup>	The Netherlands	760	Retrospective cohort
	Cooper <i>et al</i> <sup>7</sup>	UK	517	Retrospective cohort
	Khan <i>et al</i> <sup>33</sup>	Canada	2412; overlap with Perry <i>et al</i> <sup>12</sup>	Prospective cohort (secondary analysis)
	Perry <i>et al</i> <sup>20</sup>	Canada	1999; overlap with Perry <i>et al</i> <sup>12</sup>	Prospective cohort
	Perry <i>et al</i> <sup>12</sup>	Canada	3132	Prospective cohort
	Perry <i>et al</i> <sup>23</sup>	Canada	1204 had CT <6 hours	Prospective before/after
	Valle Alonso <i>et al</i> <sup>30</sup>	Spain	85	Retrospective cohort
LP	Brunell <i>et al</i> <sup>34</sup>	Sweden	453	Retrospective cohort
	Cooper <i>et al</i> <sup>7</sup>	UK	309 had LP	Retrospective cohort
	Dupont <i>et al</i> <sup>35</sup>	USA	117 had LP	Retrospective cohort
	Gangloff <i>et al</i> <sup>36</sup>	Canada	706	Retrospective cohort
	Heiser <i>et al</i> <sup>37</sup>	USA	676	Retrospective cohort
	Horstman <i>et al</i> <sup>38</sup>	The Netherlands	30	Retrospective cohort
	Migdal <i>et al</i> <sup>39</sup>	USA	245	Retrospective cohort
	Perry <i>et al</i> <sup>40</sup>	Canada	220	Prospective cohort (substudy)
	Perry <i>et al</i> <sup>41</sup>	Canada	1739	Prospective cohort (substudy)
	Sansom <i>et al</i> <sup>42</sup>	UK	60	Retrospective cohort
	Valle Alonso <i>et al</i> <sup>30</sup>	Spain	74 had LP	Retrospective cohort
CTA	Alons <i>et al</i> <sup>44</sup>	The Netherlands	70	Retrospective cohort
	Alons <i>et al</i> <sup>45</sup>	The Netherlands	88	Retrospective cohort and meta-analysis
History and examination	Locker <i>et al</i> <sup>2</sup>	UK	353	Retrospective cohort
	Perry <i>et al</i> <sup>46</sup>	Canada	747	Prospective cohort
	Backes <i>et al</i> <sup>47</sup>	The Netherlands	247	Retrospective cohort

CSF, cerebrospinal fluid; CTA, CT angiography; LP, lumbar puncture; SAH, subarachnoid haemorrhage.

### Clinical decision rules

Thirteen studies assessed the clinical decision rules developed by Perry *et al* for screening patients according to the presence of clinical characteristics associated with a high risk of SAH.<sup>13–25</sup> The predecessors of the Ottawa SAH Rule (sometimes termed the ‘Canadian clinical decision rules 1, 2 and 3’) were evaluated in six studies. Results of these studies can be found in online supplemental file 2. Rule 1 was refined to develop the final Ottawa SAH Rule, which states that alert patients with new

severe atraumatic headache, reaching maximum intensity within 1 hour, require investigation if one of the following are present: age  $\geq 40$  years, neck pain/stiffness, witnessed loss of consciousness, onset during exertion, thunderclap headache or limited neck flexion.<sup>21</sup>

A summary of the diagnostic performance of the Ottawa SAH Rule in the individual studies and pooled results generated from the bivariate meta-analysis are presented in table 2. Perry *et al* (2017) is excluded,<sup>22</sup> due to patient overlap with the larger Perry

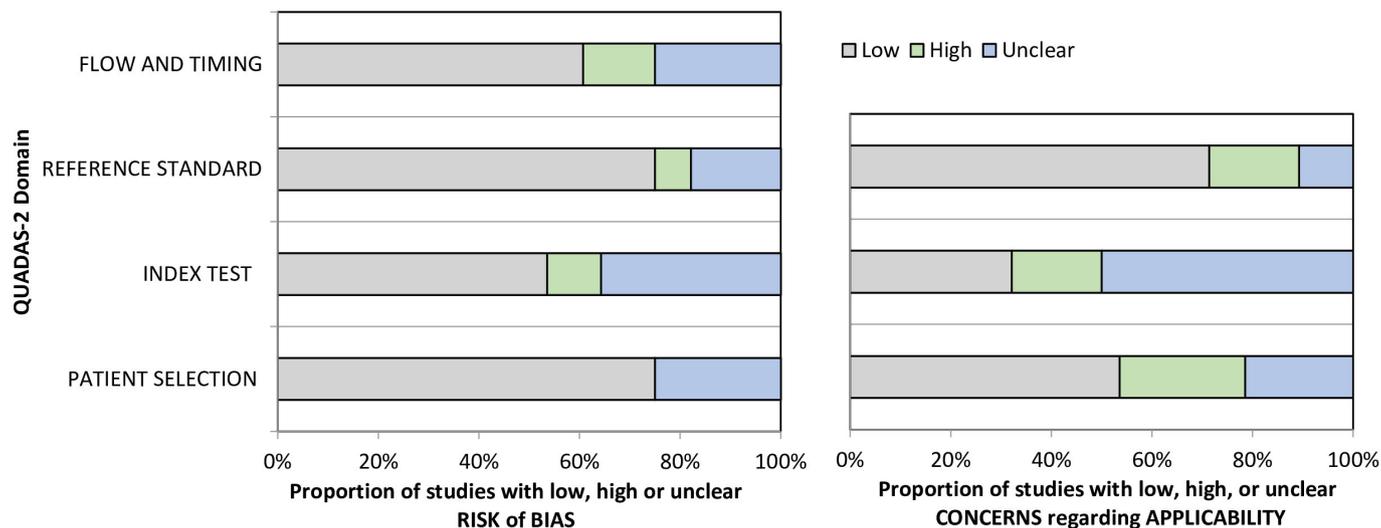


Figure 2 Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) results.

*et al* (2020) study.<sup>23</sup> The overall SAH prevalence in the studies ranged from 1.6%<sup>24</sup> to 10%<sup>14</sup> with a population-weighted mean prevalence of 5.0%. The Ottawa SAH Rule is highly sensitive, but specificity was low; strict application of the rule would result in 76% of SAH-negative patients undergoing further investigation with no additional benefit. There was considerable heterogeneity in false positive rates (FPR), potentially due to study population differences or inconsistent application of the rule. No studies assessed the accuracy of the Ottawa SAH Rule in patient subgroups by time to headache peak.

**Pathway of CT followed by LP**

The pathway of non-contrast CT followed by LP was assessed in six studies.<sup>7 26–30</sup> Only one reported complete diagnostic data, so meta-analysis was not performed. Overall, the pathway was highly sensitive, but specificity was low in some studies owing to the high FPR for LP. Importantly, this pathway also identified other significant pathologies, such as intracerebral haemorrhage, brain tumour and meningitis. More detailed results for this pathway can be found in online supplemental file 2.

**Computed tomography**

The diagnostic accuracy of CT was assessed in nine studies,<sup>7 12 20 23 26 30–33</sup> although three studies had significant patient overlap,<sup>12 20 33</sup> therefore, only the results for the largest of the three are presented.<sup>12</sup>

**CT undertaken within 6 hours of headache onset**

Four studies of CT <6 hours from headache onset were included in bivariate meta-analysis (table 3).<sup>12 23 30 32</sup> In all four studies, CT scans were assessed by neuroradiologists or radiologists who routinely interpret head CT images. Perry *et al* (2020) classed two incidental aneurysms with traumatic tap on subsequent LP as SAH, and thus as false negatives. This is inconsistent with the other included studies and with our interpretation of what constitutes a false negative. Therefore, these two patients were reclassified as true negatives.

The recruitment of patients from SAH patient databases in Backes *et al*<sup>32</sup> meant that SAH patients were over-represented in the study population (41.5%). SAH prevalence ranged from 9.2%<sup>23</sup> to 12.7%<sup>12</sup> in the other three studies, with a population-weighted average prevalence of 10.8%. Assuming that these patients are representative of those presenting to EDs in practice, the pre-test probability of SAH in patients with headache who undergo CT within 6 hours is 10.8%. Using the pooled estimate of diagnostic accuracy, the post-test probability of having suffered a SAH after a negative <6 hour CT result is 0.15%. Assuming a hypothetical follow-up test (eg, LP) has 100% accuracy, this means that 658 (95% CI 250 to 1749) patients would have to undergo further investigation to identify a single case of SAH.

One additional study assessed the diagnostic accuracy of CT <6 hours, but was excluded from the meta-analysis as it did

Study	N	Sens (%)	95% CI	Spec (%)	95% CI	FNR (%)	95% CI	FPR (%)	95% CI
Perry <i>et al</i> <sup>21</sup>	2131	100	100 to 100	15.3	13.7 to 16.8	0.0	0.0 to 0.0	84.7	83.2 to 86.3
Bellolio <i>et al</i> <sup>13</sup>	454	100	100 to 100	7.6	5.17 to 10.1	0.0	0.0 to 0.0	92.4	89.9 to 94.8
Yiangou <i>et al</i> <sup>25</sup>	162	100	100 to 100	38.7	31.4 to 46.6	0.0	0.0 to 0.0	61.0	53.4 to 68.6
Cheung <i>et al</i> <sup>14</sup>	500	94.0	87.4 to 100	32.9	28.5 to 37.2	6.0	0.0 to 12.6	67.1	62.8 to 71.5
Chu <i>et al</i> <sup>15</sup>	137	100	100 to 100	22.4	15.3 to 29.4	0.0	0.0 to 0.0	77.6	70.6 to 84.7
Pathan <i>et al</i> <sup>19</sup>	145	100	100 to 100	44.3	36.1 to 52.5	0.0	0.0 to 0.0	55.7	47.5 to 63.9
Wu <i>et al</i> <sup>24</sup>	913	100	100 to 100	37.0	33.8 to 40.1	0.0	0.0 to 0.0	63.0	59.9 to 66.2
Perry <i>et al</i> <sup>23</sup>	3672	100	100 to 100	12.7	11.6 to 13.9	0.0	0.0 to 0.0	87.3	86.1 to 88.4
Pooled (n=8)	8114	99.5	90.8 to 100	23.7	15.5 to 34.4	0.49	0.00 to 9.2	76.3	65.6 to 84.5

FNR, false negative rate; FPR, false positive rate; N, number; Sens, sensitivity; Spec, specificity.

**Table 3** Diagnostic performance of CT (<6 hours from headache onset)

Study	N	Sens (%)	95% CI	Spec (%)	95% CI	FNR (%)	95% CI	FPR (%)	95% CI
Perry <i>et al</i> <sup>12</sup>	953	100	100 to 100	100	100 to 100	0.0	0.0 to 0.0	0.0	0.0 to 0.0
Backes <i>et al</i> <sup>32</sup>	135	100	100 to 100	100	100 to 100	0.0	0.0 to 0.0	0.0	0.0 to 0.0
Valle Alonso <i>et al</i> <sup>30</sup>	85	100	100 to 100	98.7	96.1 to 100	0.0	0.0 to 0.0	1.3	0.0 to 3.9
Perry <i>et al</i> (reclassified) <sup>23</sup>	1204	97.2	94.2 to 100	100	100 to 100	2.8	0.0 to 5.8	0.0	0.0 to 0.0
Pooled (n=4)	2377	98.7	96.5 to 100	100	99.7 to 100	1.34	0.50 to 3.52	0.00	0.00 to 0.34

FNR, false negative rate; FPR, false positive rate; N, number; Sens, sensitivity; Spec, specificity.

not report sufficient diagnostic accuracy data to construct a 2×2 table to calculate sensitivity and specificity.<sup>26</sup> In this study, 760 patients had a negative CT (assessed by a staff radiologist) and subsequently underwent LP; 7% of CSF samples were initially considered positive for SAH, but subarachnoid blood was identified in only one patient on review by two neuroradiologists and a neurologist. The negative predictive value for detection of blood on CT by staff radiologists was 99.9% (95% CI 99.3 to 100).

#### CT undertaken at any time interval from headache onset

Three studies of CT undertaken at any time interval from headache onset were included in bivariate meta-analysis (table 4).<sup>7 12 32</sup> In all three studies, CT scans were assessed by neuroradiologists or radiologists who routinely interpret head CT images. The prevalence of SAH in patients undergoing CT at any time since headache onset was lower than in those who underwent CT within 6 hours. Prevalence was 2.7% in the study by Cooper *et al*<sup>7</sup> and 7.7% in the study by Perry *et al*.<sup>12</sup> As noted above, SAH patients were over-represented in the Backes *et al* study population (35.2%).<sup>32</sup>

The pooled sensitivity of CT at any time since headache onset was 94.1% (95% CI 91.0 to 96.2). This result includes patients who had CT <6 hours, as well as CT >6 hours, from symptom onset. Results from Perry *et al*<sup>12</sup> and Backes *et al*<sup>32</sup> suggest CT scans performed >6 hours after symptom onset have significantly poorer performance, reporting sensitivities of 85.7% (95% CI 78.3 to 90.9) and 90.0% (95% CI 76.3 to 97.2), respectively. The bimodal nature of the diagnostic performance of CT means that the 'CT at any time' statistics are misleading, as the timing of CT has a significant impact on the pre-test and post-test probabilities of SAH.

One additional CT study compared interpretation by emergency physicians (images viewed on standard resolution desktop screens) with the reference standard of neuroradiologists' readings (images viewed using dedicated high definition screens).<sup>31</sup> The sensitivity of CT interpreted by emergency physicians was 84% (95% CI 63.9 to 95.5) and specificity was 95% (95% CI 90.9 to 97.2). However, this study was considered to have a high risk of bias due to the difference in hardware used between the two specialties for examining CT images.

**Table 4** Diagnostic performance of CT (at any time)

Study	N	Sens (%)	95% CI	Spec (%)	95% CI	FNR (%)	95% CI	FPR (%)	95% CI
Perry <i>et al</i> <sup>12</sup>	3132	92.9	89.7 to 96.2	100	100 to 100	7.08	3.8 to 10.3	0.00	0.0 to 0.0
Backes <i>et al</i> <sup>32</sup>	247	97.6	94.4 to 100	100	100 to 100	2.38	0.0 to 5.6	0.00	0.0 to 0.0
Cooper <i>et al</i> <sup>7</sup>	510	92.9	79.4 to 100	100	100 to 100	7.14	0.0 to 20.6	0.00	0.0 to 0.0
Pooled (n=3)	3889	94.1	91.0 to 96.2	100	100 to 100	5.92	3.85 to 8.99	0.00	0.00 to 0.00

FNR, false negative rate; FPR, false positive rate; N, number; Sens, sensitivity; Spec, specificity.

#### Lumbar puncture

The diagnostic accuracy of LP in patients judged to be SAH-negative using CT was assessed in 11 studies.<sup>7 30 34-42</sup> The method of assessing CSF for xanthochromia varied, with Canadian and American studies predominantly using visual inspection and UK and European studies predominantly using spectrophotometry. LP was not always undertaken ≥12 hours from symptom onset. The standard UK NHS practice is to take the CSF sample ≥12 hours from symptom onset to allow xanthochromia to develop, with samples analysed using spectrophotometry.<sup>43</sup>

#### Spectrophotometric CSF analysis

Three studies reported diagnostic accuracy data for spectrophotometric CSF analysis following negative CT (table 5).<sup>7 36 40</sup> Samples were analysed for presence of bilirubin using the UK National External Quality Assessment Service protocol/assay.<sup>43</sup> The prevalence of SAH in these studies was only 0.65%, likely due to prescreening with CT. The FPR (and subsequent rate of angiography) was particularly high in Perry *et al* (2006), perhaps due to reported limitations in the spectrophotometric equipment used by the authors. The FPR in the more recent studies was substantially lower and likely better represents the diagnostic accuracy of CSF spectrophotometry in current practice.

Three further studies assessed CSF spectrophotometry in patients who underwent LP after negative CT, but reporting was insufficient for meta-analysis.<sup>34 38 42</sup> Horstman *et al* included 30 patients with a negative CT result for whom bilirubin was detected in the CSF; aneurysms were identified in 13 patients; however, all cases presented 4–14 days after symptom onset.<sup>38</sup> Brunell *et al* included 453 patients, 400 (88%) of whom presented with thunderclap headache; 14 (3%) patients had a pathological diagnosis based on LP, most commonly aseptic meningitis, and 5 (1.1%) had SAH.<sup>34</sup> Four of the five SAH patients had non-aneurysmal SAH which did not require surgical intervention and the other SAH patient had reduced consciousness, therefore did not strictly meet the inclusion criteria for this review.<sup>34</sup> Sansom *et al* included 60 CT-negative patients with thunderclap headache; all samples were negative for xanthochromia but 8/60 CSF examinations were abnormal for other CSF parameters (protein, glucose, cells, microscopy), with cerebral infarction confirmed in two of these patients on subsequent investigation.<sup>42</sup>

**Table 5** Diagnostic performance of spectrophotometric CSF inspection (UK National External Quality Assessment Service)

Study	N	Sens (%)	95% CI	Spec (%)	95% CI	FNR (%)	95% CI	FPR (%)	95% CI
Perry <i>et al</i> <sup>40</sup>	220	100	100 to 100	83.0	78.0 to 88.0	0.0	0.0 to 0.0	17.0	12.0 to 22.0
Gangloff <i>et al</i> <sup>36</sup>	706	100	100 to 100	98.1	96.8 to 99.1	0.0	0.0 to 0.0	1.9	0.9 to 2.9
Cooper <i>et al</i> <sup>7</sup>	309	100	100 to 100	96.8	94.8 to 98.7	0.0	0.0 to 0.0	3.3	0.1 to 5.2
Pooled (n=3)	1235	100	100 to 100	95.2	86.0 to 98.5	0.00	0.00 to 0.00	4.78	1.52 to 14.0

FNR, false negative rate; FPR, false positive rate; N, number; Sens, sensitivity; Spec, specificity.

### Visual CSF inspection

Five studies examined the diagnostic accuracy of visible xanthochromia in CT-negative patients with further investigation and follow-up used as a reference standard.<sup>35 36 39–41</sup> Three studies included sufficient information to calculate diagnostic accuracy (table 6). Sensitivity varied widely (50%–93%), due to the low prevalence of SAH (2%). The pooled false negative rate of 15% for visual inspection was higher than that for spectrophotometric analysis (0%).

Migdal *et al* assessed 245 patients with ‘low risk clinical features’, which aligned with the population in this review, but identified no cases of SAH. However, 13/245 (5.3%) patients had LP-related complications that resulted in a return visit to the ED or hospitalisation.<sup>39</sup> Perry *et al* examined the diagnostic accuracy of visible xanthochromia in ‘abnormal’ CSF samples drawn from 1739 (mostly) CT-negative patients; there were 15 (0.9%) patients classed as having aneurysmal SAH, 7 of whom had visible xanthochromia in their CSF.<sup>41</sup>

### Red blood cell-based CSF analysis thresholds

Two studies explored methods to distinguish SAH from ‘traumatic tap’, where blood enters the CSF sample due to the LP procedure itself. Perry *et al* found that the presence of fewer than  $2000 \times 10^6/L$  red blood cells (RBCs) with no xanthochromia excluded a diagnosis of aneurysmal SAH (sensitivity 100% (95% CI 74.7 to 100), specificity 91.2% (95% CI 88.6 to 93.3)) in patients who had previously undergone CT.<sup>41</sup> Heiser *et al* assessed the same RBC cut-off, reporting 81.6% sensitivity (95% CI 68.0 to 91.2) and 97.3% specificity (95% CI 95.7 to 98.4); the incidence of traumatic LP was 24.4%.<sup>37</sup> These results are not directly comparable to those reported by Perry *et al*,<sup>41</sup> as this population was not prescreened with CT.

Finally, Valle Alonso *et al* assessed 74 patients who underwent LP (method of analysis not specified) following negative CT <6 hours.<sup>30</sup> LP was positive in one patient and inconclusive in two; further imaging ruled out bleeding in all three patients. Seven patients experienced postpuncture headache, two of whom were admitted for pain control.

### CT angiography

Two small studies assessed CTA after normal CT/LP; no cases of SAH were identified, although other vascular abnormalities (including incidental aneurysms, cerebral venous thrombosis and reversible vasoconstriction syndrome) were identified.<sup>44 45</sup>

### History and examination

Three studies explored the use of historical and emergent clinical factors as predictors of SAH.<sup>2 46 47</sup> Two studies investigated the adequacy of assessment for SAH and one study assessed neurological examination for neck stiffness as a predictor of SAH. Using physicians’ clinical suspicion had a sensitivity 93% and specificity of 49%.<sup>46</sup> Presence of individual clinical factors (age >65 years, temperature >38°C, systolic BP >160 mm Hg, neck stiffness) were poor predictors of secondary headache (sensitivity 37.8%, specificity 82.1%).<sup>2</sup> Presence of neck stiffness was more strongly predictive of SAH in patients who had other high-risk clinical characteristics (eg, age ≥40 years, vomiting, transient loss of consciousness).<sup>47</sup> Recording of history in medical records was poor.<sup>2 46 47</sup>

### DISCUSSION

In summary, the Ottawa SAH Rule does little to aid clinical decision making for patients with sudden onset severe headache. The FPR was high, such that 76% of SAH-negative patients would undergo further investigation with CT and/or LP with no diagnostic value with regard to SAH, resulting in greater healthcare resource use and higher rates of adverse events related to LP and CT radiation exposure. Evidence on use of the rule in patient subgroups by time to headache peak is lacking but could be informative for clinical practice given the importance of headache incipency.

LP (with spectrophotometric CSF analysis) following negative CT was highly sensitive, although there was a 4.8% FPR. Spectrophotometry-based CSF analysis appeared to have a higher sensitivity but lower specificity than visual inspection for xanthochromia. Two studies reported rates of LP-related complications resulting in a return to the ED or hospitalisation (5%–10%). In view of the reduced sensitivity of CT >6 hours after headache onset, LP may be beneficial in these patients where a clinical suspicion of SAH remains. The CT–LP pathway also identified other significant pathologies, such as intracerebral haemorrhage, brain tumour and meningitis, meaning that its value could extend beyond the identification of SAH.

Non-contrast CT <6 hours from headache onset, with CT scans assessed by a neuroradiologist or radiologist who routinely interprets head CT images, is highly accurate for identifying SAH, and results in a very low post-test probability of SAH. This means that very large numbers of patients (estimated at 658)

**Table 6** Diagnostic performance of visual CSF inspection across identified studies

Study	N	Sens (%)	95% CI	Spec (%)	95% CI	FNR (%)	95% CI	FPR (%)	95% CI
Perry <i>et al</i> <sup>40</sup>	220	50.0	0.0 to 100	96.8	94.4 to 99.1	50.0	0.0 to 100	3.21	0.9 to 5.6
Dupont <i>et al</i> <sup>35</sup>	117	92.9	79.4 to 100	95.1	91.0 to 99.3	7.1	0.0 to 20.6	4.85	0.7 to 9.0
Gangloff <i>et al</i> <sup>36</sup>	706	80.0	44.9 to 100	98.7	97.9 to 99.5	20.0	0.0 to 55.1	1.28	0.5 to 2.1
Pooled (n=3)	1043	84.9	60.0 to 95.5	97.6	95.3 to 98.8	15.1	4.5 to 40.1	2.43	1.23 to 4.75

FNR, false negative rate; FPR, false positive rate; N, number; Sens, sensitivity; Spec, specificity.

would have to undergo further testing to yield an additional case of SAH.

However, the relatively high rate of false positive LP results (4.8% using spectrophotometry) is likely to lead to yet more testing downstream with the potential for diagnosing incidental aneurysms, leading to difficult decisions about invasive procedures. A 2016 survey of UK clinicians reported a higher risk tolerance for missed SAH diagnoses among emergency clinicians than neurospecialists, with the former accepting over 2.5 times the risk of a missed SAH (2.8% vs 1.1%;  $p=0.03$ ), and the latter more likely to advocate routine LP following a negative CT result (74% vs 39%;  $p=0.01$ ).<sup>4</sup> Emergency clinicians were also more inclined to omit LP if CT had been conducted within 6 hours of headache onset (35% vs 3%;  $p=0.002$ ).

Draft guidelines by the National Institute for Health and Care Excellence (publication delayed due to COVID-19) recommend that when there is no evidence of SAH on CT images taken <6 hours from symptom onset, LP should not be routinely offered, and alternative diagnoses should instead be considered.<sup>48</sup> However, we consider that in smaller centres without access to specialist neuroradiology expertise, or radiologists who routinely interpret head CTs, the accuracy of early CT may be reduced; studies included in our meta-analyses benefited from neuroradiology expertise. Introduction of universal access to expert interpretation of CT images could improve SAH-related patient outcomes through optimised targeting of further investigations while increasing efficiency of resource allocation. This may be achieved through widened neuro-specific training and teleradiology using other centres with relevant expertise. While interpretation of CT images using diagnostic deep learning algorithms (artificial intelligence) has the potential to improve consistency across centres, this has yet to be reliably demonstrated in high-quality studies.<sup>49</sup>

The prevalence of SAH was higher in patients who received CT <6 hours from headache onset than in the wider population of patients presenting to the ED with sudden onset severe headache (10.8% vs 7.0%). It is unclear whether this difference in pre-test probability can be assumed to exist at the point of patient assessment in the ED. Instead, triage based on severity of symptoms may have reduced wait time for CT, equally, symptom severity associated with true SAH could drive earlier presentation.

A limitation of this review was the substantial heterogeneity in the study methods and population characteristics of the included studies. The evidence base included too few patients, given the rarity of SAH events, missed diagnoses and alternative non-SAH pathologies. This led to heterogeneity in the results of some meta-analyses, and potentially meant uncertainty was underestimated in others.

There was a lack of research evidence on the small subgroup of patients who present to hospital several days after headache onset. Diagnosis of SAH in such patients is particularly challenging and there is a lack of guidance and consistency in how these patients are assessed.

## CONCLUSIONS

The Ottawa SAH Rule rules out further investigation in only a small proportion of patients; its introduction into practice could result in substantially increased rates of unnecessary investigation. Assuming the availability of neuroradiology expertise, early head CT (<6 hours) appears to be sufficient to rule out SAH in patients with sudden onset severe headache in the vast majority of patients. CT undertaken >6 hours from headache onset is much less sensitive, therefore, LP is more likely to be beneficial,

where a clinical suspicion of SAH remains. Risk tolerance of the patient and the physician, the expertise of the CT reader and consequences of additional investigations must all be considered.

**Twitter** Taj Hassan @tajekhassan

**Contributors** MW was involved in all aspects of the systematic review process including study selection, data extraction, validity assessment, synthesis of the included studies and drafting the protocol and manuscript. MW is responsible for the overall content as guarantor. RH was involved in reviewing economic studies and commented on the protocol and manuscript. AE provided expertise in evidence synthesis and project management, input at all stages of the project and commented on the protocol and manuscript. MH was responsible for devising the search strategies, carrying out the literature searches, maintaining the literature database and writing the sections of the protocol and final report relating to the searches. JS provided clinical expertise, input at all stages of the project and commented on the protocol and manuscript. TH provided clinical expertise, input at all stages of the project and commented on the protocol and manuscript. MSR provided clinical expertise, input at all stages of the project and commented on the protocol and manuscript. AH provided clinical expertise, input at all stages of the project and commented on the protocol and manuscript. John Williams provided expertise as a patient collaborator. He provided input at all stages of the project and commented on the protocol and manuscript. RW was the principal investigator who led the application for funding and took overall managerial responsibility for the project. She was responsible for the day-to-day management of the systematic review, was involved in study selection, data extraction, validity assessment, synthesis of the included studies and drafting the protocol and manuscript.

**Competing interests** All authors have completed the ICMJE uniform disclosure form at [www.icmje.org/coi\\_disclosure.pdf](http://www.icmje.org/coi_disclosure.pdf) and declare: all authors had financial support from the National Institute for Health Research (NIHR) Research for Patient Benefit (RfPB) Programme for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous 3 years; no other relationships or activities that could appear to have influenced the submitted work.

**Patient and public involvement** Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the 'Methods' section for further details.

**Patient consent for publication** Not applicable.

**Ethics approval** Not applicable.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** All data relevant to the study are included in the article or uploaded as supplementary information. Not applicable.

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution 4.0 Unported (CC BY 4.0) license, which permits others to copy, redistribute, remix, transform and build upon this work for any purpose, provided the original work is properly cited, a link to the licence is given, and indication of whether changes were made. See: <https://creativecommons.org/licenses/by/4.0/>.

## ORCID iDs

Matthew Walton <http://orcid.org/0000-0003-1932-3689>

Marc Stuart Randall <http://orcid.org/0000-0002-3196-182X>

## REFERENCES

- Goldstein JN, Camargo CA, Pelletier AJ, *et al*. Headache in United States emergency departments: demographics, work-up and frequency of pathological diagnoses. *Cephalalgia* 2006;26:684–90.
- Locker T, Mason S, Rigby A. Headache management--are we doing enough? An observational study of patients presenting with headache to the emergency department. *Emerg Med J* 2004;21:327–32.
- Dubosh NM, Bellolio MF, Rabinstein AA, *et al*. Sensitivity of early brain computed tomography to exclude aneurysmal subarachnoid hemorrhage: a systematic review and meta-analysis. *Stroke* 2016;47:750–5.

- 4 Lansley J, Selai C, Krishnan AS, *et al*. Subarachnoid haemorrhage guidelines and clinical practice: a cross-sectional study of emergency department consultants' and neurospecialists' views and risk tolerances. *BMJ Open* 2016;6:e012357.
- 5 Sayer D, Bloom B, Fernando K, *et al*. An observational study of 2,248 patients presenting with headache, suggestive of subarachnoid hemorrhage, who received lumbar punctures following normal computed tomography of the head. *Acad Emerg Med* 2015;22:1267–73.
- 6 Carpenter CR, Hussain AM, Ward MJ, *et al*. Spontaneous subarachnoid hemorrhage: a systematic review and meta-analysis describing the diagnostic accuracy of history, physical examination, imaging, and lumbar puncture with an exploration of test thresholds. *Acad Emerg Med* 2016;23:963–1003.
- 7 Cooper JG, Smith B, Hassan TB. A retrospective review of sudden onset severe headache and subarachnoid hemorrhage on the clinical decision unit: looking for a needle in a haystack? *Eur J Emerg Med* 2016;23:356–62.
- 8 McInnes MDF, Moher D, Thoms BD, *et al*. Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies: the PRISMA-DTA statement. *JAMA* 2018;319:388–96.
- 9 Whiting PF, Rutjes AWS, Westwood ME, *et al*. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011;155:529–36.
- 10 Reitsma JB, Glas AS, Rutjes AWS, *et al*. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* 2005;58:982–90.
- 11 Simmonds MC, Higgins JP. A general framework for the use of logistic regression models in meta-analysis. *Stat Methods Med Res* 2016;25:2858–77.
- 12 Perry JJ, Stiell IG, Sivilotti MLA, *et al*. Sensitivity of computed tomography performed within six hours of onset of headache for diagnosis of subarachnoid hemorrhage: prospective cohort study. *BMJ* 2011;343:d4277.
- 13 Bellolio MF, Hess EP, Gilani WI, *et al*. External validation of the Ottawa subarachnoid hemorrhage clinical decision rule in patients with acute headache. *Am J Emerg Med* 2015;33:244–9.
- 14 Cheung HY, Lui CT, Tsui KL. Validation and modification of the Ottawa subarachnoid hemorrhage rule in risk stratification of Asian Chinese patients with acute headache. *Hong Kong Med J* 2018;24:584–92.
- 15 Chu KH, Keijzers G, Furyk JS, *et al*. Applying the Ottawa subarachnoid hemorrhage rule on a cohort of emergency department patients with headache. *Eur J Emerg Med* 2018;25:e29–32.
- 16 Kelly A-M, Klim S, Edward S, *et al*. Sensitivity of proposed clinical decision rules for subarachnoid hemorrhage: an external validation study. *Emerg Med Australas* 2014;26:556–60.
- 17 MacDonald A, Sparksman D, Ramesh A. Retrospective validation of three clinical decision rules to aid the diagnosis of subarachnoid hemorrhage in patients presenting with acute headache. *Br J Neurosurg* 2012;26:137.
- 18 Matloob SA, Roach J, Marcus HJ, *et al*. Evaluation of the impact of the Canadian subarachnoid hemorrhage clinical decision rules on British practice. *Br J Neurosurg* 2013;27:603–6.
- 19 Pathan A-S, Chakarova E, Tarique A. To head CT scan or not: the clinical quandary in suspected subarachnoid hemorrhage; a validation study on Ottawa subarachnoid hemorrhage rule. *Adv J Emerg Med* 2018;2:e28.
- 20 Perry JJ, Stiell IG, Sivilotti MLA, *et al*. High risk clinical characteristics for subarachnoid hemorrhage in patients with acute headache: prospective cohort study. *BMJ* 2010;341:c5204.
- 21 Perry JJ, Stiell IG, Sivilotti MLA, *et al*. Clinical decision rules to rule out subarachnoid hemorrhage for acute headache. *JAMA* 2013;310:1248–55.
- 22 Perry JJ, Sivilotti MLA, Sutherland J, *et al*. Validation of the Ottawa subarachnoid hemorrhage rule in patients with acute headache. *CMAJ* 2017;189:E1379–85.
- 23 Perry JJ, Sivilotti MLA, Émond M, *et al*. Prospective implementation of the Ottawa subarachnoid hemorrhage rule and 6-hour computed tomography rule. *Stroke* 2020;51:424–30.
- 24 Wu W-T, Pan H-Y, Wu K-H, *et al*. The Ottawa subarachnoid hemorrhage clinical decision rule for classifying emergency department headache patients. *Am J Emerg Med* 2020;38:198–202.
- 25 Yiangou A, Nikolenko N, Noreikaite J, *et al*. Impact of subarachnoid hemorrhage Canadian clinical decision rules for investigation of acute headache: a retrospective case note review. *The Lancet* 2017;389:S103.
- 26 Blok KM, Rinkel GJE, Majoie CBLM, *et al*. Ct within 6 hours of headache onset to rule out subarachnoid hemorrhage in nonacademic hospitals. *Neurology* 2015;84:1927–32.
- 27 Dutto L, Meineri P, Melchio R, *et al*. Nontraumatic headaches in the emergency department: evaluation of a clinical pathway. *Headache* 2009;49:1174–85.
- 28 Perry JJ, Stiell I, Wells G, *et al*. Diagnostic test utilization in the emergency department for alert headache patients with possible subarachnoid hemorrhage. *CJEM* 2002;4:333–7.
- 29 Perry JJ, Spacek A, Forbes M, *et al*. Is the combination of negative computed tomography result and negative lumbar puncture result sufficient to rule out subarachnoid hemorrhage? *Ann Emerg Med* 2008;51:707–13.
- 30 Valle Alonso J, Fonseca Del Pozo FJ, Vaquero Álvarez M, *et al*. Sudden headache, lumbar puncture, and the diagnosis of subarachnoid hemorrhage in patients with a normal computed tomography scans. *Emergencias* 2018;30:50–3.
- 31 Austin R, Price J, Boyle A. Can emergency physicians accurately interpret computed tomography scans performed for suspected nontraumatic subarachnoid hemorrhage: a cross-sectional study. *Eur J Emerg Med* 2018;25:447–8.
- 32 Backes D, Rinkel GJE, Kemperman H, *et al*. Time-Dependent test characteristics of head computed tomography in patients suspected of nontraumatic subarachnoid hemorrhage. *Stroke* 2012;43:2115–9.
- 33 Khan M, Sivilotti MLA, Bullard MJ, *et al*. Factors influencing time to computed tomography in emergency department patients with suspected subarachnoid hemorrhage. *Emerg Med J* 2017;34:20–6.
- 34 Brunell A, Ridefelt P, Zelano J. Differential diagnostic yield of lumbar puncture in investigation of suspected subarachnoid hemorrhage: a retrospective study. *J Neurol* 2013;260:1631–6.
- 35 Dupont SA, Wijdicks EFM, Manno EM, *et al*. Thunderclap headache and normal computed tomographic results: value of cerebrospinal fluid analysis. *Mayo Clin Proc* 2008;83:1326–31.
- 36 Gangloff A, Nadeau L, Perry JJ, *et al*. Ruptured aneurysmal subarachnoid hemorrhage in the emergency department: clinical outcome of patients having a lumbar puncture for red blood cell count, visual and spectrophotometric xanthochromia after a negative computed tomography. *Clin Biochem* 2015;48:634–9.
- 37 Heiser H, Gimarc K, Andrews-Dickert R, *et al*. 229 external validation of a clinical prediction rule for the differentiation of traumatic lumbar punctures from aneurysmal subarachnoid hemorrhage. *Ann Emerg Med* 2015;66:S84.
- 38 Horstman P, Linn FHH, Voorbij HAM, *et al*. Chance of aneurysm in patients suspected of SAH who have a 'negative' CT scan but a 'positive' lumbar puncture. *J Neurol* 2012;259:649–52.
- 39 Migdal VL, Wu WK, Long D, *et al*. Risk-Benefit analysis of lumbar puncture to evaluate for nontraumatic subarachnoid hemorrhage in adult ED patients. *Am J Emerg Med* 2015;33:1597–601.
- 40 Perry JJ, Sivilotti MLA, Stiell IG, *et al*. Should spectrophotometry be used to identify xanthochromia in the cerebrospinal fluid of alert patients suspected of having subarachnoid hemorrhage? *Stroke* 2006;37:2467–72.
- 41 Perry JJ, Alyahya B, Sivilotti MLA, *et al*. Differentiation between traumatic TAP and aneurysmal subarachnoid hemorrhage: prospective cohort study. *BMJ* 2015;350:h568.
- 42 Sansom LT, Azad R, Datta S. Predictive value of CSF xanthochromia in CT negative thunderclap headache. *Cerebrovasc Dis* 2014;37:547.
- 43 UK National External Quality Assessment Scheme for Immunochemistry Working Group. National guidelines for analysis of cerebrospinal fluid for bilirubin in suspected subarachnoid hemorrhage. *Ann Clin Biochem* 2003;40:481–8.
- 44 Alons IME, van den Wijngaard IR, Verheul RJ, *et al*. The value of CT angiography in patients with acute severe headache. *Acta Neurol Scand* 2015;131:164–8.
- 45 Alons IME, Goudsmit BFJ, Jellema K, *et al*. Yield of computed tomography (CT) angiography in patients with acute headache, normal neurological examination, and normal non contrast CT: a meta-analysis. *J Stroke Cerebrovasc Dis* 2018;27:1077–84.
- 46 Perry JJ, Stiell IG, Wells GA, *et al*. Attitudes and judgment of emergency physicians in the management of patients with acute headache. *Acad Emerg Med* 2005;12:33–7.
- 47 Backes D, Rinkel GJE, Sturkenboom AJM, *et al*. Time-Dependent test characteristics of neck stiffness in patients suspected of nontraumatic subarachnoid hemorrhage. *J Neurol Sci* 2015;355:186–8.
- 48 National Institute for Health and Care Excellence. *Subarachnoid hemorrhage caused by a ruptured aneurysm: diagnosis and management. Draft for consultation*, 2021. Available: <https://www.nice.org.uk/guidance/GID-NG10097/documents/draft-guideline> [Accessed 15 Mar 2021].
- 49 Nagendran M, Chen Y, Lovejoy CA, *et al*. Artificial intelligence versus clinicians: systematic review of design, reporting Standards, and claims of deep learning studies. *BMJ* 2020;368:m689.