Issues in the Evaluation of Treatments for Venous Leg Ulcers

A thesis submitted in partial fulfilment of the requirements for the degree of PhD by publication

by:

Cynthia Paola Iglesias Urrutia MSc

Department of Health Sciences

THE UNIVERSITY of York

June 2006

Abstract

In this thesis I discuss a number of issues related to the evaluation of healthcare technologies for the treatment of venous leg ulcers. I begin by describing the current process of technology assessment for medical devices in wound care and point out some of the challenges healthcare decision makers face when trying to ascertain the value for money of these technologies. The main case study used in this thesis is then introduced (the VenUS I study). This is followed by a description of the evidence base concerning the clinical, cost-effectiveness and impact on health related quality of life of therapeutic medical devices in wound care, of which venous leg ulcers are an aspect. Two methods to enhance postal follow-up data collection in randomised controlled trials are then proposed and tested in two opportunistic studies. Previously identified methodological shortcomings associated with existing evidence of the clinical and cost-effectiveness of wound care interventions are addressed in the context of the largest randomised controlled trial of two compression systems for the treatment of venous leg ulcers. Robust statistical methods that acknowledge relevant characteristics associated with the primary clinical and economic measures of health benefit and costs are used to analyse data from VenUS I. The potential contribution of a recently proposed methodological framework based on Bayesian decision analytic modelling and Bayesian value of information analysis is investigated. Finally I present a discussion of the contribution of this body of work, and its impact on research activity, and conclude making some recommendations for future research.

Contents

Abstract ii
Contents iii
Tables and Figures v
Acknowledgements vi
Relevant publications and declaration of candidate's contribution vii
Abbreviations ix
Introduction 1
Section I. 7
1 General overview of the area of VLUs 7
1.1 Chronic leg ulcers 7
1.2 Epidemiology 7
1.3 Aetiology of venous leg ulcers 8
1.4 Treatment 8
1.4.1 Effectiveness of compression therapy (prior to VenUS I) 9
1.4.2 Effectiveness of other treatments (up to date) 10
1.5 Health related quality of life (HRQoL) 10
1.5.1 HRQoL from a clinical and economic perspective 11
1.5.2 The impact of VLU on HRQoL (prior to VenUS I) 13
1.6 Economics of VLUs 14
1.6.1 Economic evaluation 14
1.6.2 Economic evaluations in VLUs 15
Section II 17
2 Measuring the impact of VLU on HRQoL 17
2.1 Background 17
2.2 Methods 20
2.2.1 Opportunistic Study I (paper 1 in Appendix A) 20
2.2.2 Opportunistic Study II (paper 2 in Appendix A) 21
2.2.3 VenUS I HRQoL data (paper 3 in Appendix A) 21

2.3 Results 22
2.4 Conclusions 23
Section III. 26
3 Trial-based evaluation 26
3.1 Background 26
3.2 Methods 27
3.3 Results 29
3.3.1 Clinical Analysis 29
3.3.2 Economic Evaluations 30
3.4 Conclusion 31
Section IV 33
4 Research prioritisation in VLU 33
4.1 Background 33
4.2 Methods 34
4.3 Results 35
4.4 Conclusions 36
Section V 38
5 Concluding remarks and overall recommendations 38
5.1 Contributions to knowledge 38
5.1.1 Optimisation of data collection 38
5.1.2 Measurement of HRQoL in people with VLUs 39
5.1.3 Clinical and cost-effectiveness of high compression for VLUs 41
5.1.4 Research prioritisation in VLUs 43
5.2 Recommendations for future research 45
Appendix A 48
Appendix B 49
Appendix C 53
Bibliography 57

Tables and Figures

Table 1. Treatments for VLUs included in *Clinical Evidence* 12

Table 2. Evidence on VLUs HRQoL post VenUS I 40

Table 3. Summary of contributions to knowledge 44

Figure 1. Alternative primary aims of Health Technology Assessments. 6

Figure 2. Quorum chart of economic evidence. 56

Acknowledgements

Ad Andrea e al nostro bimbo nella pancia, la mia inspirazione e forza...

I would like to give most praise to David Torgerson and Nicky Cullum for having made real a dream, for their continuous support and guidance, and above all for their invaluable friendship.

A special thanks to Karl Claxton for embarking with me in this project as co-supervisor and for helping me find my way when confusion seemed to be winning.

To Michael Drummond and Trevor Sheldon for their advice and generosity, and for sharing not only their academic knowledge but also their personal wisdom.

To Christine Godfrey, Pat Ansell and Mark Sculpher for their insightful comments on the integrative chapter of this thesis.

To Andrea Manca and Christian Asseburg for sharing their knowledge on quantitative data analysis.

I am most grateful to Sarah Collins for proof reading this manuscript but also for her support and encouragement throughout the completion of this thesis.

Como siempre a Yolanda, Guillermo C, Guillermo, Lupita, Nydia y Ulises por su cariño y comprensión.

A mis amigas Carmina, Leticia y Maret por su entusiasmo y aliento.

Finally, to all the people who accompanied me in this journey and made it a treasured life experience (Javier, Meehee, Gigi , Cesar, Atzhiri, Fabio, Hege, Adriana, Werner, Dawn, Helen, Rocio, Steven, Seokyung, Deborah, Pedro, Astrid, Juan, Cristina, Joaquin).

Relevant publications and declaration of candidate's contribution

First publication: Iglesias C, Torgerson D. Does length of questionnaire matter? A randomised trial of response rates to a mailed questionnaire. *J Health Serv Res Policy* 2000;5(4):219-221.

The candidate participated in the design of the study, was responsible for data collection, conducted all statistical analyses, and was responsible for the preparation and submission of the manuscript and subsequent revisions.

Second publication: Iglesias CP, Birks YF, Torgerson DJ. Improving the measurement of quality of life in older people: the York SF-12. *Q J Med* 2001:94:695-698.

The candidate conceived the research question, proposed and developed an alternative layout for the SF-12 questionnaire, participated in the design of the study, was responsible for data collection, conducted all statistical analyses, prepared and submitted the manuscript and subsequent revisions.

Third publication: Iglesias CP, Birks Y, Nelson EA, Scalon E, Cullum NA. Quality of life of people with venous leg ulcers: a comparison of the discriminative and responsive characteristics of two generic and a disease specific instruments. *Quality of Life Research* 2005;14 (7); 1705 – 1718.

The candidate participated in the implementation of the study, was the data manager, statistician and health economist for the study, prepared the analysis

plan and conducted all statistical analyses of the quality of life data, prepared and submitted the manuscript and subsequent revisions.

Fourth publication: Nelson EA, Iglesias CP, Cullum N, Torgerson DJ. Randomised clinical trial of 4-layer and short stretch compression bandages for venous leg ulcers (VenUS I). *British Journal of Surgery* 2004; 91: 1292-1299.

The candidate participated in the implementation of the study, was the data manager, statistician and health economist for the study, prepared the analysis plan and conducted all statistical analyses of the clinical data, and contributed to the first and subsequent drafts of the manuscript.

Fifth publication: Iglesias CP, Nelson EA, Cullum N, Torgerson DJ. Economic analysis of VenUS I: a randomised trial of two bandages for treating venous leg ulcers. *British Journal of Surgery* 2004; 91: 1300-13006.

The candidate participated in the implementation of the study, was the data manager, statistician and health economist for the study, prepared the protocol and conducted all statistical analyses of the economic data, prepared and submitted the manuscript and subsequent revisions.

Sixth publication: Iglesias CP, Claxton K. Comprehensive Decision-Analytic Model and Bayesian Value-of-Information Analysis : Pentoxifylline in the Treatment of Chronic Venous Leg Ulcers. Pharmacoeconomics 2006;24(5):465-78.

The candidate conceived the research question, reviewed the literature, conducted the data extraction and statistical analyses, prepared and submitted the manuscript and subsequent revisions.

Abbreviations

Venous Leg Ulcer(s)	VLU(s)
Health Related Quality of Life	HRQoL
United Kingdom	UK
National Health Service	NHS
Research and Development	R&D
Health Technology Assessment	HTA
Economic Evaluation(s)	EE
Cost-Effectiveness Analysis	CEA
Cosy-Utility Analysis	CUA
Cost-Benefit Analysis	CBA
Cost-Consequences Analysis	CCA
Quality Adjusted Life Years	QALYs
Euroqol 5 dimensions	EQ-5D
General Practice	GP

Introduction

Publicly funded healthcare systems are under increasing financial and political pressure. Policy makers are expected to optimise the allocation of their limited healthcare resources in a way that maximises society's health benefit and ensures a fair distribution of the benefits. To this end, several jurisdictions have set up formal healthcare evaluation processes to establish whether new technologies are good value for money.[1-3] In England and Wales the National Institute for Health and Clinical Excellence (NICE) has responsibility for issuing recommendations regarding the adoption of healthcare technologies, informed by evidence regarding their clinical and cost-effectiveness.[4] The availability of high quality evidence is fundamental for the successful completion of this process.

This thesis discusses a number of issues associated with the assessment of medical devices for wound care specifically indicated for venous leg ulcers (VLUs). Unlike pharmaceuticals, for which a stringent evaluation process is in place, the introduction of therapeutic medical devices into the European market requires only evidence of safety and technical performance.[5] This limited requirement for evidence of therapeutic benefit results in the introduction of a myriad of products with similar indications and unproven differential effectiveness. This is true with respect to wound care products, most of which are licensed and marketed with relatively weak clinical and cost-effectiveness evidence.

The limited incentives to demonstrate clinical and cost-effectiveness of medical devices provided by the current licensing system has resulted in the widespread use of case studies and case series as the evidential basis for marketing. Such is the

case for all wound care products including those indicated for VLU treatment. However, observational studies are widely regarded as providing lower quality evidence of clinical effectiveness than well conducted randomised controlled trials (RCTs).[6] Whilst the number of therapeutic RCTs in VLUs has increased during the last decade, most of them still have important methodological limitations.[7-9] These limitations are detailed in Appendix B – to describe them briefly here, they include: i) inappropriate comparisons; ii) small samples unsupported by *a priori* sample size calculations; iii) inadequate duration of follow-up ; iv) between group imbalances in prognostic factors at baseline; v) use of intermediate and sub-optimal outcome measures (for a more detailed description of these shortcomings, see Appendix B).

Most economic evidence on the cost-effectiveness of VLU treatments is accrued from RCTs, and consequently the validity of any trial-based evaluation is entirely dependent on the quality of the RCTs.[10] Similarly, the validity of the results of model based economic evaluations depends on the quality of the evidence used to populate the model and the suitability of its structure.[11;12] The generally weak evidence base in VLUs has therefore resulted in severe limitations associated with existing trial and model-based economic evaluations. Moreover, the methodological quality of these evaluations does not adhere to existing guidelines.[13] Frequently encountered shortcomings include: i) use of uninformative intermediate outcomes; ii) estimates of cost only based on unit cost of treatments; iii) use of medians rather than means to estimate average health benefits and costs (for a detailed discussion see Appendix B).

It can be argued that in general the poor methodological quality of the clinical and cost-effectiveness evidence for medical devices generates an additional hurdle for the assessment of these technologies. The Health Technology Assessment (HTA)

process has been defined as "*the evaluation of the properties, effectiveness and the direct and indirect impacts of health technologies.*"[14] In the absence of a fixed set of activities to be performed as part of an HTA process, activities are determined according to the primary aim of the evaluation which may vary and include an assessment of: the effectiveness of a health technology in specific populations; its cost; any intended or unintended consequences of a health technology for individuals, families, healthcare system, and society.[14] A graphical representation of HTA processes with a range of different objectives is described in Figure 1. The HTA processes in both the UK National Health Service HTA Programme (UK NHS HTA) and NICE primarily focus on the clinical and cost-effectiveness of health technologies from the perspective of the UK NHS and Personal Social Services.[4;13] As represented in Figure 1, the HTA process is an iterative approach with different primary objectives which are intrinsically interconnected.[14;15]

The overall aim of the research described in the published papers discussed in this thesis is to produce high quality research evidence to inform / facilitate HTA in VLUs. Specifically, this research concentrates on two aspects of HTA: the assessment of the clinical and cost-effectiveness of health technologies (single activities involved in this process with a note of those informed by our findings are listed in Figure 1) and methodological issues in primary research. To achieve my overall aim I pursued the following objectives: i) to improve the measurement of effectiveness in VLU using final and composite outcomes that can also be used to measure health benefit from an economic perspective (e.g. QALYs); ii) to optimise the collection of primary outcome data using postal surveys in elderly populations; iii) to rigorously assess the clinical and cost-effectiveness of high compression systems for VLU; and iv) to test the

potential contribution of a recently proposed framework for the evaluation of healthcare technologies.

In 1998, the UK NHS HTA programme commissioned a multicentre pragmatic RCT comparing the effectiveness of two high compression bandage systems for the treatment of VLUs (the VenUS I Study). The question addressed by this trial had been generated by a Cochrane review.[16] VenUS I constitutes the main research base for the development of this thesis (see Section 3.1 for a detailed description).

Interestingly, following the commissioning of the VenUS I study, a number of industry-funded RCTs on high compression systems for VLUs were also conducted in the UK and Europe.[17-22] These concurrent RCTs had significantly smaller sample sizes and follow-up periods than the VenUS I study, and hence their findings were published whilst VenUS I was underway. In order to preserve a sense of the chronology of events and in the iterative spirit of the HTA process, I discuss the clinical (Section 1.4.1) and economic evidence (Section 1.6.2.1) on high compression for VLUs at two points in time, prior to the conception and funding of the VenUS I study in 1998 (Section 1.4.1) and from that point up to January 2006 (Section 5.1.3).

VenUS I was also the first study in which preference-based HRQoL scores were used to estimate health benefit in an economic analysis. During its conduct, however, a number of cross sectional and interventional studies investigated the impact of VLUs from clinical and economic perspectives using personal interviews and disease specific and generic health related quality of life (HRQOL) instruments.[23-25] The results of VenUS I are set in the context of this emerging body of evidence in Section 5.1.2 and the evidence prior to VenUS is discussed in Section 1.5.2.

This integrative chapter is divided into five sections. Section I provides a general overview of the state of the art in the area of VLUs. Section II addresses the topic of HRQoL: its measurement using disease specific and generic instruments; the impact of VLUs; and its use as a measure of effectiveness in RCTs (papers 1,2 and 3 in Appendix A). Section III is dedicated to the analysis of the clinical and cost-effectiveness of two alternative systems of high compression for VLUs (papers 4 and 5 in Appendix A). The application of a recently proposed framework for the evaluation of healthcare technologies to the area of VLUs is tested in Section IV (paper 6 in Appendix A). Finally, a discussion of the main findings, limitations of the work and its contribution to knowledge are the subject of Section V.

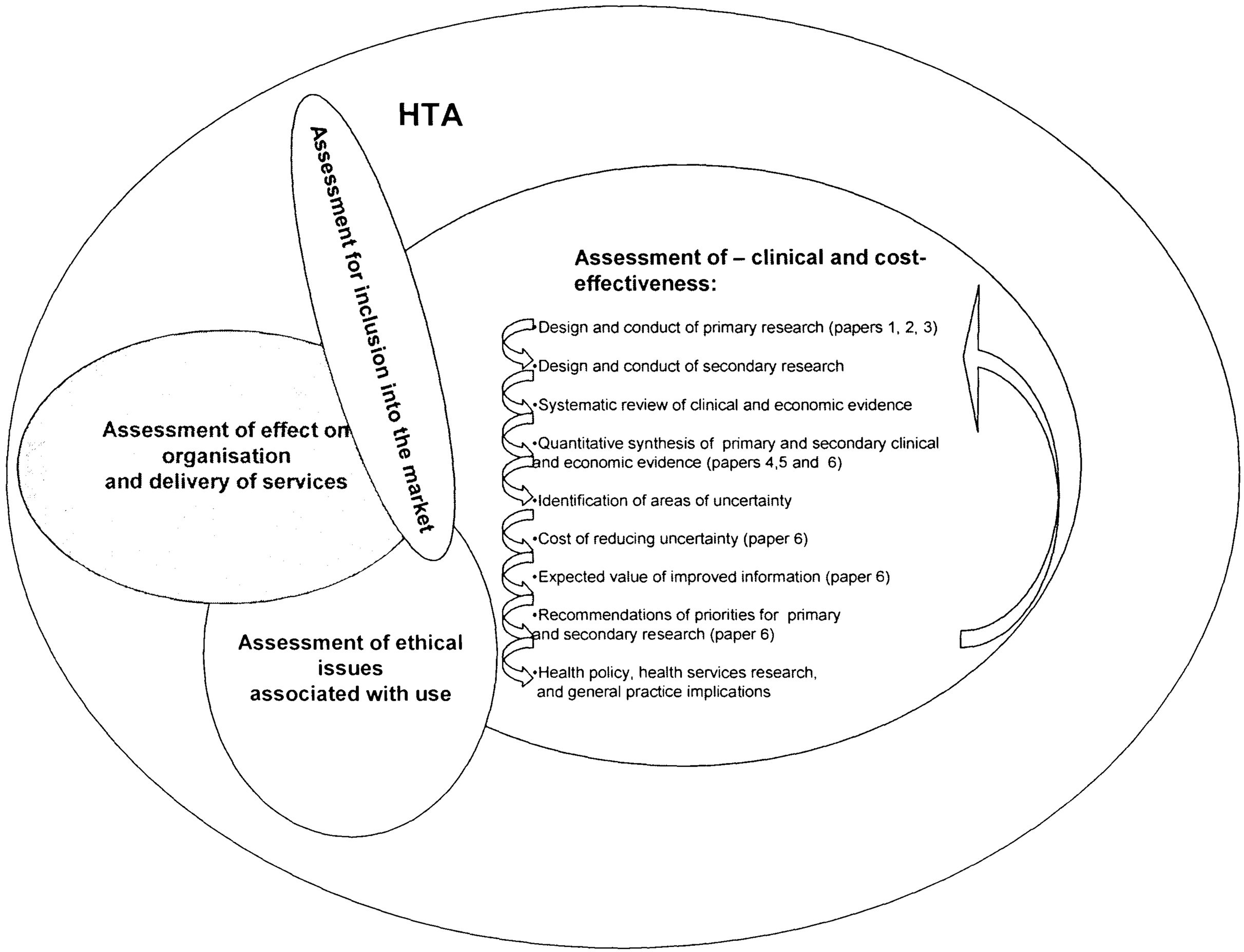


Figure 1. Alternative primary aims of Health Technology Assessments.

Section I.

1 General overview of the area of VLUs

1.1 Chronic leg ulcers

Chronic leg ulcers have been defined as... "*a loss of skin below the knee on the leg or foot that takes more than six weeks to heal*".[26] The chronicity of the disease is determined by an underlying health condition, frequently venous insufficiency, that causes the skin to breakdown by some unknown mechanism. Effective treatments for venous insufficiency are not well developed, hence the chronic nature of the sequelae, including ulceration.

1.2 Epidemiology

Leg ulcers are mainly present in ageing populations, with women being more frequently affected than men.[27] The extent of leg ulceration in the UK has been investigated in a number of epidemiological studies which suggest that the prevalence of this condition in the British adult population is between 0.045% to 1.8%.[28;29] The largest UK based study suggests that approximately 25% of the identified chronic leg ulcers were active (i.e. open) at any point in time.[30;31]. Uncomplicated venous disease has been identified as the main underlying cause of leg ulceration with venous disease being apparent in 43% to 81% of leg ulcers.[27;29;32;32;33] VLUs are typically a recurring condition.

1.3 Aetiology of venous leg ulcers

The aetiology of VLUs is only partially understood. Deep vein thrombosis (DVT), which results in damage to the deep veins leading to venous insufficiency, has been identified as a risk factor for the development of VLUs.[34] The cellular mechanisms involved in the breakdown of the skin of the leg, however, have not been fully explained.[35] Browse and Burnand found evidence that venous insufficiency was associated with back flow of blood and an increased hydrostatic pressure in the superficial veins.[36] It is in an attempt to reduce the hydrostatic pressure in the veins and to aid venous return that external compression therapy is applied to the ulcerated leg.

1.4 Treatment

The first line treatment of VLUs comprises the application of wound dressings and external compression (standard wound care).[37;37] The application of external compression through bandages, hosiery (stockings, socks, tights) or combinations of these, is also used to prevent further skin breakdown.[38] Standard wound care treatment can be complemented by use of pharmacological drugs, topical agents, debriding agents, laser treatment, therapeutic ultrasound, skin grafting, and venous surgery.[39] Prognosis and the patient's initial response to treatment should determine departures from the "standard treatment". For example, VLUs that fail to heal whilst in compression may require an adjunct therapy, such as pentoxifylline, to expedite the healing process.[40]

When all the alternative methods for compression and wound dressings indicated for VLUs in the British National Formulary (BNF)[41] are taken into account, there are approximately 130,000 potentially different clinical treatment pathways for VLUs. This may unnecessarily complicate healthcare decision making in VLUs. In Table 1, I have described all the interventions indicated for the treatment of VLUs in the UK, their indication, the available RCT evidence and potential different classes of products, as described in the *BNF* and *Clinical Evidence*.[38;41] The latter is a BMJ internet based publication, the mission of which is to "*…summarise the current state of knowledge and uncertainty about the prevention and treatment of clinical conditions, based on thorough searches and appraisal of the literature. […] It describes the best available evidence from systematic reviews, RCTs and observational studies where appropriate,…*"[38]

The clinical effectiveness evidence in this field is generally weak for the reasons outlined in the introduction and described in Appendix B.

1.4.1 Effectiveness of compression therapy (prior to VenUS I)

In 1997 the evidence base for external compression in the treatment of VLUs was synthesised in a Cochrane systematic review [16], which indicated that, in the absence of arterial disease: i) compression was more effective than no compression; ii) high compression was more effective than low compression; iii) there was insufficient evidence of a differential effect associated with different high compression systems (multi-layered, single layered, elastomeric, non-elastomeric). To address these issues the review concluded by recommending the conduct of well designed RCTs accompanied by economic evaluation analyses.

1.4.2 Effectiveness of other treatments (up to date)

Wound dressings are typically applied directly to the open ulcer underneath compression bandages. The superiority of any wound dressing over another has not been demonstrated.[7;38] Similarly, a paucity of evidence on the effectiveness of other therapies indicated for the treatment of VLUs has been identified (see Table 1). Evidence of the clinical effectiveness of pentoxifylline, a vasodilator drug, as an adjunct to compression in the treatment of VLUs, was identified in a systematic review.[40] Indirect evidence of the effectiveness of compression hosiery to prevent recurrent VLUs was provided in two RCTs of compression stockings for the prevention of VLUs.[42;43]. In the view of the methodological weaknesses associated with these studies their findings should be used / interpreted with caution.

1.5 Health related quality of life (HRQoL)

In 1947 the World Health Organisation defined health not only as the absence of disease or infirmity but as "a state of complete physical, mental and social well-being."[44] Ever since, "health related quality of life" (HRQoL) has become a frequently used umbrella concept that encompasses a number of different dimensions / domains of health. To date there is no universally accepted definition of HRQoL, although a range of descriptions are available in the literature. It has been argued that existing definitions vary depending on both, "*the aims and perspective of the research and the instruments that are developed and applied to attain such aims.*"[45] In health services research, measurements of HRQoL are often made from a clinical or an economic perspective.

1.5.1 HRQoL from a clinical and economic perspective

This concept has been used from a clinical perspective to explore the impact of a health condition on specific dimensions of health, such as: daily routine, physical, social, intellectual, emotional, psychological functioning, economic status, health status, well-being or life satisfaction, among others. Such research has used both quantitative and qualitative approaches. In health economics, the concept of HRQoL has been closely associated with "utility",[10] an economic term that describes the strength of preference of individuals, in this case, preference for a health outcome or health status under conditions of uncertainty, relative to two health states: perfect health and death. The more preferable a health status or health outcome, the greater the utility expected to be associated with it. Unlike clinical / psychometric measurements of HRQoL, "utility measurements" allow comparisons of the HRQoL impact associated with a range of different health outcomes, hence their relevance for HTA.

In a clinical context, measurement of HRQoL is usually conducted in a certainty framework through the administration of psychometric instruments often using scaling methods such as rating scales, category scaling, visual analogue scales or ratio scales. In economics, measurement methods based on "choice under uncertainty" (i.e. standard gamble) are the only ones deemed to measure utility.[46] Difficulties with the practical implementation of standard gamble make it a less popular alternative than choice methods under certainty such as time trade off, paired

Health Technology	Indication	Evidence[c]	Categories
External Compression[a]			
Bandages	Primary treatment for venous leg ulcers (vlu)	A (28 RCTs)	17
Stockings	Mainly prevention but can be used for treatment of vlu	A (2 RCTs)	3
Intermittent pneumatic compression	Adjuvant to compression to accelerate healing	B (4 RCTs)	1
Wound Dressings[b]			
Simple low adherent	Wound contact layer to use on granulating tissue	A (15 RCTs)	14
Foam	Wound contact layer to use on hypergranulating tissue or highly exuding wounds	B (2 RCTs)	8
Film	Secondary dressing for mildly exuding wounds	B (2 RCTs)	2
Hyaluronic acid –derived	Rehydration and debridement of dry, sloughy or necrotic wounds	B (1 RCT)	2
Alginates (semi-occlusive)	Moderately or heavily exuding wounds	D (1 RCT)	11
Hydrocolloid (occlusive)	Rehydratation and debridement of dry, sloughy or necrotic wounds	AA (9 RCTs)	16
Pharmacological Treatments			
Pentoxifylline	Promote healing	A (11 RCTs)	2
Oral flavonoids	Promote healing	C (2 RCTs)	
Oral sulodexide	Promote healing	C (2 RCTs)	
Oral aspirin	Promote healing	C (1 RCT)	4
Oral rutosides	Promote healing	D (2 RCTs)	1
Oral thromboxane α_2 antagonist	Promote healing	D (1 RCT)	
Oral zinc	Promote healing	D (5 RCT)s	
Systemic mesoglycan	Promote healing	C(1 RCT)	1
Granulocyte-macrophage	Promotion of cellular functions to stimulate wound healing	C (1 RCT)	3
Topical agents			
Antimicrobials	Infected ulcers	C (14 RCTs)	4
Cultured allogenic bilayer skin	Induction of tissue regeneration	C (1 RCT)	
Calcitonin plus polypeptide	Promote healing	D (1 RCT)	
Mesoglycan	Promote healing	D (1 RCT)	
Negative pressure	Promote healing	D (1 RCT)	
Recombinant keratinocyte growth factor 2	Promote healing	D (1 RCT)	
Autologous platelet lysate	Promote healing	D (1 RCT)	
Debriding therapies			
Wound dressing	Remove slough and necrotic tissue from wound	D (23 RCTs)	17
Larval therapy	Remove slough and necrotic tissue from wound	C (1 RCT)	2
Surgical debridement	Remove slough and necrotic tissue from wound	E	
Other treatments			
Skin Grafting	Induction of tissue regeneration	C (6 RCTs)	
Superficial vein surgery	Prevent future ulceration	B (4 RCTs)	4
Ultrasound	Promote healing	C(7 RCTs)	
Laser (low level)	Promote healing	C (7 RCTs)	

Table 1. Treatments for VLUs included in *Clinical Evidence*

[a] External compression main characteristics: i) Level of compression (Light (14-17 mmHg), Moderate (18-24 mmHg), High (25 to - 35 mmHg), Extra high (up to 60 mmHg)); ii) Elasticity (elastomeric, non-elastomeric); iii) Layers of compression (single, multi-layer); iv) Delivery method (bandage, stockings, rigid cast, pneumatic).

[b] Variations in dressings are associated with the following characteristics: I) Adherence; ii) Moisture; iii) Retention; iv) Permeability; v) Sterilisation; vi) Allergenicity; vii)Absorbency

[c] Summary of evidence from Clinical Evidence: A = evidence of benefit; AA = Evidence of no benefit; B = Conflicting evidence, C = Insufficient evidence; D = No evidence of benefit; and E = No evidence.

comparison, and person trade off.[11] While a correlation between clinical and economic measurements of HRQoL is apparent, an analytically explicit description of the association might not be known. Thus it is relevant to consider both types of measurements in the context of HTA, as discussed further in Section 2.

1.5.2 The impact of VLU on HRQoL (prior to VenUS I)

By 1998 three disease specific instruments had been claimed as suitable for the measurement of the HRQoL of individuals with VLUs. The 34-item Hyland is the only instrument that has specifically been developed to measure the HRQoL of individuals with a VLU.[47] The chronic lower limb venous insufficiency questionnaire (CIVIQ)[48] and the Freiburger questionnaire of quality of life in venous disease (FLQA)[49] aim to measure HRQoL in people with venous insufficiency who do not necessarily have an ulcer. The CIVIQ is a 20-item scale which measures four dimensions of HRQoL (psychological, physical, and social functioning and pain) with good reliability and validity (face, content and construct). The CIVIQ is also reported to be sensitive to changes in venous insufficiency status after two months.[48] The FLQA comprises 83 items within seven scales (physical complaints, everyday life, social life, emotional status, therapy, satisfaction, occupation) and was associated with good validity and sensitivity to changes in venous insufficiency status after three months. To date, however, an English translation of this instrument is not publicly available.

Prior to VenUS I, HRQoL in people with and without leg ulcers had been reported in two studies. Flett et al found that leg ulceration was associated with greater problems with activity and mobility, more pain, lower levels of self-esteem, and higher levels of

negative affect.[50] Roe et al observed poorer scores for pain, energy, life satisfaction and depression reported by individuals with leg ulcers.[51]

1.6 Economics of VLUs

VLUs are extremely costly; annual treatment costs to the UK NHS were estimated at between £230 and £650 million, at 1990-1991 prices.[52]

1.6.1 Economic evaluation

Economic evaluation analyses of health technologies have been recognised as a tool to inform the decision making process in healthcare.[4] In the context of HTA, economic evaluation has been defined as "*the comparative analysis of competing alternative courses of action in terms of both the costs and benefits associated with them.*"[10] In full economic evaluation the estimates of incremental cost and incremental health benefit are combined in a single measure of cost-effectiveness, such as the incremental cost-effectiveness ratio or the incremental net benefit.[10;53]

There are three different analytical approaches used in *full* economic evaluations of healthcare technologies. The main differences between these arise from the way in which health benefits associated with the alternative treatments under comparison are measured. In cost-effectiveness analyses (CEA), health benefits are measured using natural units (e.g. ulcer free days, absolute and relative reductions in ulcer size). Evaluations in which health benefits are measured using a composite measure that weights life expectancy with a preference-based utility score, such as EQ-5D, to estimate quality adjusted life years are known as cost-utility analyses (CUA). The

ability of CUA to compare interventions for different indications has made it the preferred method of evaluation by NICE.[13] In cost-benefit analysis (CBA), health benefits are measured in monetary terms; recently this has been done using discrete choice methods.[54]

Cost-consequences analyses (CCA) are *partial* economic evaluations in which the health benefits and costs associated with the alternatives are not combined in a single measure of cost-effectiveness.[10]

1.6.2 Economic evaluations in VLUs

The economic evaluations in the area of VLUs discussed in this section were identified in a systematic review conducted in 2004 as a basis for an update of the Royal College of Nursing Guidelines for the Management of VLUs.[39] The electronic search strategy used, as well as a QUORUM flowchart of the evidence identified (Figure 2), are described in Appendix C.[55]

In general, the validity of the economic evaluations identified is compromised by methodological shortcomings associated with the clinical studies on which they are based and by the absence of rigour in the measurement of overall treatment cost and health benefits.

1.6.2.1 Cost-effectiveness of compression systems for VLU (prior to VenUS I)

As recently as 1998, the cost-effectiveness of high compression systems for VLU had only been explored in one trial-based cost-consequences study[56] which claimed

Unna's boot (a device used mainly in the USA) was associated with higher healing rates and less costs than hydrogel dressing.

1.6.2.2 Cost-effectiveness of other treatments for VLU (up to date)

In spite of the paucity of evidence on the clinical superiority of any wound dressing (Section 1.4.2), several studies have claimed the cost-effectiveness of hydrocolloids relative to saline gauze based on reductions in weekly nursing contacts.[57-60] Opportunistic CCAs have been conducted in an attempt to justify both the cost-effectiveness of larval therapy to remove sloughy / necrotic tissue (debride);[61] and the cost-effectiveness of managing VLU patients in co-ordinated community VLU clinics using multi-layer high compression versus community care.[62-65] Evidence of the value for money of compression stockings and an educational programme versus "do-nothing" to prevent future VLUs was found in a model based cost-utility analysis.[66] As mentioned in the introduction the evidence from these studies is associated with the shortcomings discussed in Appendix B, and hence their findings should be used / interpreted with caution.

Section II

2 Measuring the impact of VLU on HRQoL

2.1 Background

The Hyland questionnaire is the only disease specific instrument that claims to be able to measure HRQoL in individuals with VLUs.[47] As mentioned in Section 1.5.2, other disease specific HRQoL instruments used in this population focus primarily on venous insufficiency (a sequela of which is venous ulceration). The Hyland, while constructed using sound psychometric methodology, was developed in a relatively small sample of 50 people with VLUs. The VenUS I study provided an ideal opportunity to validate and further investigate the properties of this instrument in a larger population of individuals with VLUs, prior to recommending its use in treatment evaluations.

A comprehensive investigation of HRQoL in any given population as well as the use of HRQoL as an outcome measure within therapeutic RCTs often requires both generic and disease specific instruments. However, the inclusion of multiple HRQoL instruments in one survey can result in lengthy questionnaires with often poor response rates.[67] Similarly, postal surveys have been associated with poorer response rates than telephone or personal interviews.[68;69] In spite of this, postal surveys are often the only feasible method for collecting follow-up clinical and economic data in RCTs. Using any other survey methodology (e.g. personal or telephone interviews) within RCTs may be inefficient since it would require substantial human and financial resources. Furthermore, it has been argued that the quality of

responses to postal surveys could be enhanced by ensuring the face validity (readability) of HRQoL instruments.[67]

During the implementation of the VenUS I trial a small pilot study was conducted to verify the readability of a HRQoL booklet to be used as a data collection tool. The booklet initially contained the VLU specific Hyland questionnaire, two generic instruments (EQ-5D and SF-12) and the McGill pain questionnaire.[70-72] The SF-12 and the EQ-5D are both well known generic HRQoL instruments that allow the investigation of HRQoL from a clinical and economic perspective, respectively. Over a period of 2 consecutive days, one of the VenUS I Research Nurses asked all her scheduled VLU patients to complete these four HRQoL instruments.

Eight consecutive individuals (mean age 74 years) with VLU agreed to participate. As an observer during these consultations I made written comments about the interviewees' individual reactions and their comments about the booklet. All the participants stated that they found the survey took longer than they initially expected. In addition, it was noted that the layout of the SF-12 was confusing. Three of the eight interviewees missed or only partially responded to questions with a "steam and leaf" format in the SF-12. This suggested the readability of specific items within a questionnaire could differ according to variations in the question layout and the specific characteristics of the population of interest.

Based on the results from the eight initial interviews, it was considered that potential future respondents may find it difficult to complete this long and complex questionnaire. As a result, the McGill pain questionnaire was removed from the booklet and the original layout of the SF-12 modified by the candidate into twelve single closed questions; see paper 2 in Appendix A. This new booklet was re-tested

in the three older participants who had had difficulties with the previous one and 8 new participants. The quality of the responses to this second version of the HRQoL booklet improved. The three participants who previously failed to provide adequate responses now provided complete responses to the second version of the booklet.

This small pilot study suggested a shortened questionnaire with a simplified layout might be more acceptable to participants in the VenUS I study. It was important, however, to investigate these findings in larger studies as it was not possible to test, statistically, the reliability of the new measure or its expected overall completion rates on such a small sample. Large RCTs in populations similar in age to that of the participants in the VenUS I trial were necessary to formally investigate both: i) the reliability of this alternative-form (layout) of the SF-12, its overall response rates and quality of responses; and ii) whether longer questionnaires did adversely affect response rates in older participants.

To optimise the collection and measurement of HRQoL data from individuals with VLUs in postal surveys conducted as part of RCTs, the following research questions were tested:

a) Does the length of a questionnaire affect the rates of overall and partial responses to health related surveys? [73] (see paper 1 in Appendix A)

b) Does a simpler layout for the SF-12 have a positive impact on the overall and partial responses to this instrument?[74] (see paper 2 in Appendix A)

c) Is the Hyland a reliable and valid instrument to measure VLU specific HRQoL?[75] (see paper 3 in Appendix A)

d) What are the properties of the Hyland, SF-12 and EQ-5D in measuring HRQoL in individuals with VLUs?[75] (see paper 3 in Appendix A)

e) What is the impact of VLUs on HRQoL from both a clinical and an economic perspective in a UK population?[75] (see paper 3 in Appendix A)

2.2 Methods

The first two hypotheses were tested in opportunistic (piggy backed) studies in populations similar in age to that of the participants in the VenUS I trial (women aged 70 and over) conducted during the recruitment stage of a large RCT of Hip Protectors for the prevention of fractures.[76]

2.2.1 Opportunistic Study I (paper 1 in Appendix A)

A quasi-randomised trial was conducted in women age 70 years and over recruited from a general practice (GP) in North Yorkshire, England, to investigate whether the inclusion of multiple HRQoL instruments, on top of a clinical questionnaire of osteoporotic risk factors, was detrimental to the overall response rates of postal surveys within RCTs.

Four different survey booklets were compared. The first contained the clinical questionnaire; the second the EQ-5D and clinical questionnaire; the third contained the clinical questionnaire, EQ-5D and SF-12; and the fourth booklet was the same as third but the SF-12 had the alternative layout described in Section 2.1. A table describing the way in which the "steam and leaf" questions in the SF-12 were modified is provided in paper 2 in Appendix A.[74]

A total of 856 questionnaires were sent to a GP in Yorkshire where staff labelled and posted the questionnaires to women potentially eligible to participate in the Hip Protector Study.[76] Alternation was used to produce equivalent groups since there is evidence that, if strictly adhered to, alternation is as effective in preventing selection bias as random allocation.[77] In this instance we can be assured that alternation was adhered to, as the opaque envelopes were delivered to the practice staff and they had no knowledge that the survey was also testing questionnaire length.

2.2.2 Opportunistic Study II (paper 2 in Appendix A)

The second quasi-randomised trial investigated whether the two different layouts of the SF-12 had a differential effect on overall response rates in a large sample. One thousand five hundred questionnaires were mailed out; 750 questionnaires with the original layout of the SF-12 and the EQ-5D; and 750 with the EQ-5D and an alternative layout of the SF-12. Alternation was used again to produce equivalent groups. Sealed envelopes were sent to the participating UK GPs, where they were labelled and posted.

2.2.3 VenUS I HRQoL data (paper 3 in Appendix A)

Formal investigation of the third and fourth aims required data from a population with VLUs. Data from VenUS I were used to investigate the validity of the Hyland questionnaire and the discriminative and responsive characteristics of the EQ-5D, SF-12 and Hyland HRQoL instruments (see Section 3.1 for further details of the VenUS I study).

2.3 Results

The results from the first pilot study suggested that shorter postal surveys are more likely to have higher overall response rates. Increasing the length of a questionnaire from five to seven pages reduced overall response rates from 49% to 40% (an absolute difference of 9%, 95% CI 0.3% to 16.6%) in a population of women aged 70 years and over. The findings from the second pilot study indicated that while similar overall response rates were associated with the two layouts of the SF-12 questionnaire, partial responses were statistically significantly reduced with the simplified layout of the SF-12.

Baseline HRQoL data from the 387 participants in the VenUS I study suggested a two-factor solution for the third section of the Hyland. Factor one related to practical limitations, and factor two to emotional perceptions. Both factors were negatively affected among individuals. At baseline, detriments in the practical and emotional components of 47% and 36% were observed.

No statistically significant differences in Hyland (practical and emotional), EQ-5D and SF-12 (physical and mental) scores between bandage groups were observed, hence the evaluative and responsive characteristics of these instruments were investigated combining the data from both groups. The Hyland questionnaire was associated with small and moderate ability to discriminate individuals according to age, mobility, initial ulcer size and ulcer duration.

The SF-12 and EQ-5D had good evaluative properties; both instruments were sensitive to changes in HRQoL after ulcer healing at 3 months. During the one-year follow-up, improved HRQoL was reported quarterly by individuals whose VLU had

healed in the EQ-5D and SF-12. The difference in quarterly EQ-5D scores between healed and unhealed individuals was statistically significant in all four quarterly measurements. The conclusion of an episode of VLU was associated with an improvement relative to baseline scores on the mental component on the SF-12 and stabilisation of the physical component. EQ-5D scores suggested a statistically significant improvement in HRQoL associated with healing relative to active ulceration. High levels of bodily pain were reported in the SF-12 questionnaire, whilst only minor ulcer related discomfort was reported in the Hyland.

2.4 Conclusions

The importance of ensuring a thorough consideration of patient data collection instruments to be used in postal surveys within RCTs was confirmed by the results of the two opportunistic studies described here. The first study provided evidence that lengthy questionnaires and those with poor readability (face validity) are likely to have poorer overall response rates and a higher proportion of partial responses (missing items). A systematic review of best practice to design and use patient and staff questionnaires in health services research acknowledged a paucity of evidence on the impact of questionnaire appearance and design on response rates.[67] Out of 11 identified studies on the effect of questionnaire length only two were on health-related topics.[78;79] Evidence of a negative relationship between length of questionnaire and response rates was provided by these studies albeit non-statistically significant.

The second opportunistic study indicated that readability can play an important role in maximising the quality of responses to postal surveys. In an elderly population, we found that shorter questionnaires with single questions may be more appropriate than

lengthy ones with complex question structures. To the best of my knowledge this is the first study looking at alternative layouts for the SF-12, and until further research is conducted to investigate whether this is an age related issue, we would recommend the use of a simpler layout for the SF-12 in postal surveys.

The marginal contribution of the 34-item Hyland questionnaire to improve the understanding and measurement of the impact of VLUs on HRQoL is equivocal. Whilst one single VLU pain specific question from the Hyland provided valuable evidence to suggest that VLU related pain may be considerably lower than that suggested in previous studies,[50;51] its inclusion almost tripled the overall length of the HRQoL booklet used in VenUS I. Furthermore, in the context of the pilot study described in Section 2.1, participants raised concerns regarding the appropriateness, subsequent to the closure of their VLU, of a large proportion of VLU specific items in Hyland, i.e. they questioned the responsiveness of this instrument to VLU healing. Hence, the research team decided to administer Hyland only to people with VLUs.

Whist over time only minor VLU specific pain was reported in the Hyland, high levels of bodily pain were reported in the SF-12 by both healed and unhealed individuals. This may be partially explained by the fact that VLUs are more frequent among the elderly who are likely to have comorbidities. Reported bodily pain may be a consequence of these concomitant chronic health conditions. This issue needs to be further investigated in future RCTs in VLUs.

Whilst no statistically significant differences in HRQoL scores between bandage groups were observed, quarterly HRQoL scores from the EQ-5D and SF-12 provided indirect evidence of improved HRQoL associated with compression systems associated with a faster rate of healing. Statistically significant differences in EQ-5D

scores between healed individuals and people with unhealed ulcers were observed, suggesting that treatments associated with a shorter time to VLU healing will also be associated with improved HRQoL. A resolution of VLU was associated with improvements in the mental component of the SF-12 and EQ-5D.

In conclusion, simple and brief generic HRQoL instruments, SF-12 and EQ-5D, provide a comprehensive description of the impact of active ulceration in HRQoL from a clinical and economic perspective. Until the responsiveness to healing of the Hyland questionnaire is established, the use of this instrument in future RCTs in VLUs is not recommended. Furthermore, preference-based measurements from the EQ-5D allowed exploration not only of the cost-effectiveness of the two compression systems compared in VenUS I relative to each other (see Section 3), but potentially with an ample spectrum of healthcare treatments (see Section 5.2).

Section III.

3 Trial-based evaluation

3.1 Background

In 1997 the UK NHS HTA programme commissioned the VenUS I Leg Ulcer study, the largest multi-centre pragmatic RCT in the area of VLUs to date. This study compared two high compression systems in the treatment of VLU in terms of clinical and cost-effectiveness, a comparison identified as a research priority in an earlier systematic review.[16]

The VenUS I trial compared the *4-layer* bandage, regarded by many in the UK as the gold standard compression therapy for VLU, and the *short-stretch* bandage, a compression therapy frequently used in the European continent and Australia.[80] While both compression systems apply a high degree of compression to the leg (40mmHg), considerable differences are associated with their required frequency of administration and purchasing cost. In terms of usage and cost, the *short-stretch* is a reusable bandage claimed to be able to remain in place for a week with a unit cost of £5;[41] whereas the *4-layer* bandage can be used only once, can remain in place for one week and has a unit cost of £8.[41] Differences in the unit cost and wearability of these compression systems suggested that potentially substantial savings in bandage expenditure could be achieved, if the *short-stretch* bandage were shown to be of equivalent or greater effectiveness than the *4-layer* bandage.

The main objectives of the VenUS I study were to:

a) investigate the clinical effectiveness of the *short-stretch* bandage compared with *4-layer* bandage in the treatment of VLUs in the UK. [81] (see paper 4 in Appendix A)

b) determine the cost-effectiveness of *short-stretch* relative to *4-layer* in the treatment of VLUs in UK practice. [82] (see paper 5 in Appendix A)

c) explore the impact of VLUs on HRQoL from both a clinical and an economic perspective in a UK population.[75] (see paper 3 in Appendix A)

This section will focus on objectives I and II. The third objective was discussed in Section II.

3.2 Methods

The VenUS I study was conceived and conducted within an academic institution according to rigorous methodological standards.[83] Individuals with a VLU of at least: 1 week duration and 1 cm in length or width, were eligible for inclusion. Length of follow-up was one year. Central telephone randomisation was used to ensure unbiased allocation of participants to trial treatments. Participants were stratified according to ulcer area (smaller or greater than 5 cm^2); current ulcer duration (shorter or longer than 6 months); and centre, initially four (Cumbria, Leeds, West London, North Yorkshire) with other centres (Falkirk, Calderdale, East London, Newmarket, Southport) recruited later.

The primary outcome for the clinical analysis was time to healing of all VLUs on the trial leg. Median time to ulcer healing in both bandage groups was calculated using Kaplan Meier estimates of the cumulative probability of healing.[84] The log–rank test was used the investigate differences between bandage groups in the Kaplan Meier curves for time to healing and time to ulcer recurrence.[85;86] The hazard of healing was adjusted by baseline covariates in a Cox proportional hazards model, where interactions between the treatment effect and baseline covariates were also investigated.[87]

Two economic evaluations were conducted from the perspective of the UK NHS and Personal Social Services. Ulcer free days (Kaplan Meier estimates of mean time to ulcer healing) and QALYs, were the measures of health benefit in a CEA and a CUA, respectively. Quarterly utility scores reported on EQ-5D were used to compute QALYs as the area under the curve defined by these values. QALYs were adjusted by imbalances between bandage groups in the EQ-5D scores at baseline using linear multiple regression.[88]

Total volume of four categories of resource use (nurse visits, doctor visits, hospital visits and bandages used) was collected prospectively from all the participants in the VenUS I study to estimate total treatment costs. To account for the censored nature of cost and HRQoL data, the Lin method was used to estimate the mean total treatment cost per bandage group and QALYs.[89]

Non–parametric bootstrapping was used to estimate 95% confidence intervals for the mean differences in ulcer free days, QALYs and costs between bandage groups. The scenario approach to sensitivity analysis was implemented to explore the robustness of the results to variations in both: i) monthly utilisation rate of *short-stretch* bandages;

and ii) method of acquisition *4-layer* bandage (as a kit rather than as single items). The decision uncertainty associated with the relative cost-effectiveness of *short-stretch* compared with *4-layer* bandage was explored in both an incremental cost-effectiveness plane and a cost-effectiveness acceptability curve (CEAC).[90-93]

3.3 Results

Three hundred and eighty seven people from 9 centres in the UK were recruited into the VenUS I study. 192 were allocated to the *4-layer* and 192 to the *short-stretch*. Individuals in both bandage groups were similar with respect to patient characteristics (age, height, weight), ulcer characteristics (area, duration, episodes, among others) and stratification variables (area, duration, centre).

Clinical, economic and HRQoL measures of effectiveness indicated that relative to *short-stretch*, the *4-layer* was associated with a higher instantaneous probability (i.e. risk) of healing, shorter time to healing, smaller treatment cost and slightly better HRQoL.

3.3.1 Clinical Analysis

The full clinical analysis of the VenUS I study is reported in paper 4 in Appendix A.

The clinical unadjusted analysis found a non-statistically significant difference in Kaplan Meier median times to healing between bandage groups; 92 days (*4-layer* group) versus 126 (*short-stretch* group). However, when the hazard of healing was adjusted by prognostic factors in a Cox regression model, the *short-stretch* was associated with a statistically significantly lower probability of healing than *4-layer*

bandage (hazard ratio 0·72, 95% CI 0.57 to 0.91). The same model suggested a statistically significant interaction between bandage treatment and centre. The most heterogeneous results were associated with data from the smaller participating centres. In an analysis considering only data from the (larger) initial four centres participating in the VenUS I study, heterogeneity in the treatment effect by centre was no longer observed, i.e. the coefficient of the treatment centre interaction was no longer statistically significant.

Kaplan Meier cumulative probabilities of recurrence were 0.36 and 0.39 in the *4-layer* and *short-stretch* groups, respectively. A log-rank test of the Kaplan Meier curves of time to ulcer recurrence in both groups indicated no statistically significant differences between these curves.

3.3.2 Economic Evaluations

The two economic evaluation analyses are reported in full in paper 5 in Appendix A. Patients in the *4-layer* group healed on average 10.9 days before those in the *short-stretch* group (95% CI: -6.76 days to 29.06 days). Difference in mean QALYs between compression systems was -0.02 (95% CI: -0.08 to 0.04). The *4-layer* bandage cost on average £227.30 less per patient per year than the *short-stretch* bandage (95% CI: £16.53 to £448.30).

These results were robust to feasible variations in the monthly utilisation rate of *short-stretch* and the unit cost of the *4-layer*. The CEAC suggested that considerable uncertainty was associated with the decision to consider *short-stretch* as a cost effective intervention relative to *4-layer* for a large range of λ-values [£0 to £30,000],

i.e. willingness to pay values for an extra ulcer free day. *Short-stretch* was associated with approximately a 20% probability of being cost-effective.

3.4 Conclusion

This trial suggested that, in the UK, the *4-layer* bandage is more likely to be clinically and cost-effective than the *short-stretch* bandage as the first line treatment for VLU.

The economic evaluations in VenUS I indicated that despite initial expectations of cost savings associated with the *short-stretch* bandage, the *4-layer* bandage was a "dominant" alternative, i.e. *4-layer* was associated with greater health benefits and lower overall treatment costs. Initial expectations were based on the differential in unit cost between the *short-stretch* and the *4-layer* bandages and the ability of the former to be re-used. The main cost driver in the treatment of VLUs, however, is nursing time. Hence, "potential savings" are unlikely to translate into real ones, unless reductions in nursing time are generated.

In spite of claims from the manufacturers regarding the ability of *short-stretch* to remain in place for a week, the trial reported a higher frequency of nurse visits for patients in this group. Bandage slippage and poor rates of bandage reuse reported in the *short-stretch* group may account for this. In addition, VLUs are more frequent among elderly populations. The mean age of the participants in the VenUS I study was 71.6 (SD 13.21), and older participants reported having experienced difficulties washing their *short-stretch* bandages.

Evidence of the ability of the *short-stretch* bandage to remain in place for a week has been provided by research conducted in Austria and the Netherlands, where the

short-stretch bandage is considered the standard for treatment of VLUs.[19] This suggests that nurses' competence in bandage application might have an influential role in the effectiveness of external compression systems. While more research may be recommended in this area, it is important to note that, in the UK, VLU care is provided in both at dedicated clinics and hospitals, as well as within the community, where patients are usually looked after by a pool of nurses rather than a single nurse. These intrinsic characteristics of the VLU service in the UK render unfeasible a rigorous evaluation of "competence in bandage application". Nonetheless, proxy measures of nurses' competence in bandage application, such as total number of patients previously treated by a nurse using a specific bandage, may be used in future analyses where random effects multilevel models can be used to explore these variations in nurse's performance.[94]

Evidence of a potential interaction between bandage treatment and centre was identified in a Cox proportional hazards model. Future therapeutic RCTs in VLUs should acknowledge a potential for heterogeneity in the treatment effect associated with factors such as nurses' competence with bandage treatment; nurses' prior beliefs about the effects of different treatments; the degree of variation in nurses' wound care experience, and patients' concordance with bandage treatment.

The VenUS I study provides evidence of the feasibility of conducting rigorous clinical and economic evaluations of therapeutic treatments for VLU. This high quality evidence can, in turn, be used to inform the evaluation of other health technologies using decision analytic modelling as discussed in Section IV.

Section IV.

4 Research prioritisation in VLU

4.1 Background

Data from well conducted RCTs is considered the "gold standard" scientific evidence that should be used to inform treatment decisions in healthcare practice. However, in view of the large number of available alternative treatments indicated for VLU (Table1) it would be both unreasonable and inefficient to defer the evaluation of the relative clinical and cost-effectiveness of these technologies until sound RCT evidence is available.[95] The scarcity of research and development (R&D) resources, requires decision makers to counter their need for further information by substantiating their decision on existing information, and thus to free resources which can be allocated to other areas of priority (i.e. take into consideration the opportunity cost of using R&D resources in any specific area and not another).[96]

A framework based on comprehensive decision analytic modelling, Bayesian methods for evidence synthesis, and Bayesian value of perfect information (VOI) analysis has recently been proposed for the evaluation of healthcare technologies.[97] This framework aims to inform two key stages in the process of deciding whether to recommend the adoption of a health technology: i) whether a health technology should be considered clinically and cost-effective given the existing evidence base; and ii) whether a request for further information should be made given the level of decision uncertainty associated with the current evidence base.

The main objectives of the research described here were to investigate the potential of this analytic framework to:

a) assist the formal evaluation of the clinical and cost-effectiveness of alternative treatments for VLU based on existing information;[98] (paper 6 in Appendix A)

b) guide the process of prioritisation of the research agenda in VLUs identifying those areas where primary research is more likely to reduce the decision uncertainty associated with the cost-effectiveness of competing treatments.[98] (paper 6 in Appendix A)

4.2 Methods

A pharmacological adjunct therapy to external compression was used as a case study. Pentoxifylline, a drug that helps blood flow, has been proposed as a potential agent to accelerate the healing process of VLUs.[99] A systematic review of RCTs of pentoxifylline for the treatment of VLUs identified evidence of an additional beneficial effect to external compression, and a potential beneficial effect in the absence of compression, associated with pentoxifylline.[40] The poor methodological standards of RCTs on pentoxifylline, however, raised concerns regarding the validity of the findings from this review. Indeed, a number of relatively small RCTs on pentoxifylline have recently been conducted.[100;101]

The purpose of this work is to investigate whether the decision uncertainty associated with the clinical and cost-effectiveness of pentoxifylline prior to commissioning the largest confirmatory trial was likely to justify the expected cost of primary research. A comprehensive Bayesian decision model[102] was constructed to investigate the

cost-effectiveness of pentoxifylline as an adjunct to compression. The analysis was conducted at two different points in time: i) prior to the publication of the findings from the largest RCT on pentoxifylline in the presence of external compression (*prior analysis*) and ii) following the publication of the results from this study (*posterior analysis*).

Potential heterogeneity between RCTs was accounted for in the model using a Bayesian random baseline meta-analysis[103] implemented in the freely available specialist software WinBugs.[104] The model was populated using information from both previously published data [26;105-107] and data from the VenUS I trial.[80] Preference-based utility weights reported by participants in the VenUS I study with and without VLUs were used to estimate expected QALYs.[75]

A non-parametric method was used to estimate the expected VOI[108] associated with the decision to recommend pentoxifylline as a cost effective adjunct to compression for a large range of willingness to pay values for an extra QALY [£0 - £5000].

4.3 Results

The *prior* and *posterior* analyses suggested pentoxifylline was a dominant therapy, that is, relative to placebo, pentoxifylline was associated with greater health benefits and was less costly. In the *prior* analysis, patients in the pentoxifylline group on average healed 8.28 weeks (95% credibility interval (CrI) 1.89 to 14.56) earlier than patients in the placebo group. Similarly, an average 0.02 gain in QALYs (95% CrI -0.12 to 0.17), and an average reduction of £153.4 (95% CrI -£53.11 to £354.9) in overall cost was observed in the pentoxifylline group relative to the placebo group.

Estimates of the uncertainty surrounding the cost-effectiveness of pentoxifylline and the VOI in both the *prior* and *posterior* analyses suggested that further primary research on the cost-effectiveness of pentoxifylline was unlikely to be cost-effective. In the *prior* analysis, the estimated VOI for willingness-to-pay values of £0, £100, and £500 per QALY were £128,200, £127,100 and £126,700, respectively. Comparing these values with the cost of the largest compression study in VLUs (the VenUS I trial), approximately £380,000[109], suggests that at the time when the Dale trial was funded the cost of primary research was higher than the maximum expected benefit, i.e. the reduction in the decision uncertainty associated with the acquisition of additional evidence.

Incorporation of the information from the largest RCT on pentoxifylline did improve the precision of the estimate of the clinical effect associated with this drug; however this improvement was not large enough to reverse either the decision regarding the dominance of pentoxifylline or the maximum value associated with further research.

4.4 Conclusions

This analysis indicated that the cost of conducting further primary research on the effectiveness of pentoxifylline as an adjunct to compression in the treatment of VLUs is unlikely to be offset by the expected reduction in the decision uncertainty on the cost-effectiveness of this drug associated with the acquisition of additional evidence. Indeed, the evidence base considered in this analysis suggested pentoxifylline is associated with at least a 90% probability of being cost-effective for a range of values [£0 - £500] decision makers may be willing to pay for an additional QALY.

This case study illustrates the huge potential of comprehensive decision analysis, Bayesian evidence synthesis methods and Bayesian VOI analysis as a framework to evaluate medical devices for the treatment of VLUs, and as a tool to inform the research prioritisation process in this area. Had the results from this analysis been available before the largest RCT on pentoxifylline was funded, a more efficient allocation of research and development resources could have been made. To promote a more efficient allocation of scarce R&D resources in both the public and private sectors, future recommendations for primary research in the area of VLUs should be carefully evaluated, explicitly accounting for the cost opportunity of investing in an area and not another.

Finally, recently proposed quantitative methods to synthesise evidence[110;111], not only from studies with the same design but from a range of study types to explore the clinical and cost-effectiveness of all the relevant alternatives for the treatment of a health condition, should be explored in the area of VLUs.[112-114]

Section V.

5 Concluding remarks and overall recommendations

5.1 Contributions to knowledge

This thesis has explored a number of relevant issues in the assessment of therapeutic treatments for VLUs.

5.1.1 Optimisation of data collection

The optimisation of clinical and economic primary data collection using postal surveys conducted within RCTs was explored in two opportunistic studies (papers 1 and 2 in Appendix A). The first analysis suggested that lengthy surveys and those with complex layouts may prove unnecessarily challenging.[73] As a result an alternative layout for the widely used generic HRQoL questionnaire SF-12 was proposed and found to be more suitable for the populations studied and potentially others too.[74] To the best of my knowledge this is the first analysis proposing an alternative layout for the SF-12 questionnaire. Whilst there is no consensus about the effect of questionnaire length on response rates, a recently conducted systematic review and meta-analysis of RCTs evaluating the effect of questionnaire length on response rates coincides with our findings that in the context of postal surveys within RCTs shorter questionnaires should be preferred.[115]

The applicability of the findings from our two opportunistic studies to other age groups could not be investigated. Such analysis would require a large survey conducted in a sample representative of the whole of the UK adult population. The two surveys discussed here, however, were piggy-backed studies conducted in the recruitment stages of a large RCT of hip protectors.[76] Nonetheless, in a recently published large population survey using the SF-12 in its standard layout, age was a significant predictor of missing items.[116] This suggests that the findings from our two opportunistic studies may be relevant to the sector of the population in which poor data collection is more likely to be observed.

Our findings have not only informed the design of two new large RCTs in the area of VLUs recently funded by the UK NHS HTA, the Larval Therapy Study (VenUs II)[117], and the Ultrasound Study (VenUS III) [118], but also the simplified layout of the SF-12 is now used in preference to the original layout in all trials within the York Trials Unit.

5.1.2 Measurement of HRQoL in people with VLUs

A validation analysis of the only VLU specific HRQoL questionnaire available to date (Hyland) indicated that the contribution of this instrument to improve understanding of the impact on HRQoL associated with the presence of an open VLU is marginal. We found the SF-12 and EQ-5D to be associated with good responsiveness to VLU healing, thus suggesting that generic instruments may be more suitable to measure changes in HRQoL associated with the presence / absence of a VLU.

In the area of VLUs, the VenUS I study was the first one in which preference-based HRQoL measurements were used to estimate health benefit in an economic analysis.

During its conduct, a number of studies measuring the impact of open VLUs on HRQoL were published, see Table 2.[23-25]

Author	Area	Instruments	Population	Conclusion
Walters 1999[23]	UK	Medical outcomes study short form 36 (MOS SF-36); Euroqol (EQ-5D); Short form McGill pain questionnaire (SF-MPQ); Frenchay activity index (FAI)	233 people with venous leg ulcers	The 4 instruments had small to moderate discriminative abilities. The SF-MPQ had the best responsiveness to healing at 3 months and the SF-36 and EQ-5D were only able to detect changes in healing status after 12 months
Wissing 2001[24]	Sweden	Philadelphia geriatric centre multilevel assessment instrument	144 people with and without leg ulcers	VLUs were associated with significantly lower mean score for physical health, activities of daily living, cognition, time use / social behaviour, personal judgement and environmental quality
Franks 2001[25]	UK	Nottingham health profile (NHP)	383 people with venous leg ulcers	NHP had a good internal consistency, small to moderate ability to discriminate individuals according to age, mobility, initial ulcer size and ulcer duration. Unlike the SF-36 the NHP was sensitive (bodily pain dimension) to changes in healing status at 3 months

Table 2. Evidence on VLUs HRQoL post VenUS I

Two core discrepancies between these studies and the HRQoL analysis from VenUS I are distinguished. Firstly, the identification of the high levels of chronic pain reported by people with VLUs as primarily ulcer related pain in the studies published before VenUS I. Secondly, sole consideration in these studies of open VLUs as the main / only health condition negatively affecting HRQoL, thus ignoring the role of concomitant chronic health conditions. Conversely, our findings coincide with those from a study of HRQoL in people with varicose veins, a disorder encompassed in the venous insufficiency syndrome. Using a recently proposed disease specific instrument for venous disease, (the VEINES-QoL scale) Lamping et al concluded that

the HRQoL of individuals with varicose veins can only be reliably interpreted when concomitant venous disease is taken into account.[119] It is worth noting that the VEINES-QoL increases the number of disease specific instruments potentially suitable for VLUs to four, see Section 1.5.2.

VLUs are a manifestation of chronic venous insufficiency, a syndrome encompassing a spectrum of disorders ranging from varicose veins through to extensive ulceration of the leg. Future studies aiming to describe and / or measure the impact on HRQoL associated with VLUs should acknowledge this by distinguishing the dimensions of HRQoL affected by an open ulcer, detriments in HRQoL associated with venous insufficiency and those mainly related with other chronic concomitant health conditions.

Recent methodological developments in the estimation of preference-based measures from generic psychometric instruments have produced a model to estimate preference-based scores from the SF-12.[120] The level of agreement / disagreement between the preference-based scores derived from the EQ-5D and SF-12 in the VenUS I study should be investigated in a future study.

5.1.3 Clinical and cost-effectiveness of high compression for VLUs

The VenUS I study provided evidence of the "dominance" of the *4-layer* bandage relative to the *short-stretch* bandage. On average *4-layer* was associated with a shorter time to healing of VLUs, marginally better HRQoL and less overall treatment costs. Evidence of a differential bandage treatment effect by centre was identified in

the analysis of the clinical data from the VenUS I study. Formal investigation of this result using random effect multilevel models for survival data, i.e. frailty models [121], was prevented by the unavailability of data regarding factors such as nurses' competence in bandaging, prior beliefs, and preferences about different systems of external high compression. This limitation associated with the VenUS I study could be addressed at the design stage of future RCTs in VLUs.

To date, in addition to the economic evaluations in VenUS I, the clinical and cost-effectiveness of external compression in the treatment of VLUs has been investigated in six partial[17;18;22;122-124] and one full economic evaluation.[125] The cost-effectiveness of the *4-layer* system was investigated in six of these studies the findings of which coincide with those from the VenUS I study, in that, on average *4-layer* compression systems are associated with a higher proportion of patients healed and a reduction in the number of nurse visits required for VLU treatment.

In 1998, when the VenUS I study began, HRQoL had not been used as an outcome measure within therapeutic RCTs in VLUs. The advantages of using composite health outcomes that incorporate a quality of life adjustment in the evaluation of healthcare technologies have been well documented.[10] The use of quality-adjusted outcomes (e.g. QALYs), for instance, allows the direct comparison of health technologies indicated for different health conditions, thus enabling the comparison in terms of cost per QALY of, for example, the value for money of investing in cancer drugs with that of investing in anti-retroviral drugs. Future economic evaluation studies in VLUs should measure health benefit associated with the alternatives under comparison using such quality adjusted composite outcomes. This recommendation has already informed the design of the VenUS II and VenUS III studies in which QALYs have been chosen as

the primary outcome for the economic analyses to be conducted alongside these RCTs.[117;118]

5.1.4 Research prioritisation in VLUs

The potential contribution of comprehensive decision analytic modelling, Bayesian evidence synthesis methods and Bayesian value of information analysis to the formal evaluation of therapeutic treatments for VLUs was tested in a case study. To the best of my knowledge this is the first application of these methods in the area of VLUs. These analytical tools, however, are increasingly being used to inform the decision making process at NICE.[126] This illustrates the ability of this framework to inform the decision making processes regarding both the relative cost-effectiveness of treatments given their existing evidence base and whether to commission further primary research to reduce the decision uncertainty associated with the adoption of treatments.

Our analysis indicated that, given the current evidence base on pentoxifylline, there is a high probability that it is a cost-effective adjunct to high compression. Consequently, additional primary research on the cost-effectiveness of pentoxifylline is unlikely to be an efficient use of R&D resources. Pentoxifylline, while still officially unlicensed for this indication, should be considered as a valuable alternative for the treatment of those individuals whose VLU fails to heal with external compression.[40] However, the point at which an adjunct to external compression as pentoxifylline should be sought may be further investigated using decision analysis.

This VOI analysis was implemented using RCT evidence only. In the last few years, however, powerful and versatile Bayesian statistical models for evidence synthesis

have been proposed in the literature to integrate coherently all the available evidence on the relative effectiveness of health technologies using data not only from RCTs but also from clinical studies with different designs, i.e. observational studies.[111;112] This potential shortcoming associated with this analysis will be addressed in a future study in which multiple parameter evidence synthesis and networks of evidence will be used to synthesise all the existing evidence on pentoxifylline.[112]

In summary, the work published as part of this thesis represents an important contribution to wound care research and to health services research in general. The contributions made are summarised in Table 3 and supported by evidence of the direct impact of this body of work on the research activities of our research team. Furthermore, our findings have informed and / or provided the basis for a number of successful research applications: i) NHS HTA funded VenUS II and VenUS III studies; ii) MRC Special Training Fellowship in Health Services Research and Health of the Public, in which I shall explore the use of Bayesian evidence synthesis methods for the regulation of therapeutic medical devices; iii) MRC funded Vacuum Assisted Compression (VAC) study, an analysis of the VOI associated with VAC in the treatment of pressure ulcers which includes a feasibility pilot study of a potential future RCT of VAC in the same patient population.

Study	Contribution to knowledge
Questionnaire design	Improvements to the layout of a widely used quality of life instrument, which will lead to more reliable data collection in future, both in wound care and non-wound care trials and surveys.
HRQoL in venous leg ulceration	The Hyland instrument is inappropriate to investigate treatment related effects within RCTs in VLUs.
Compression bandaging trial	Demonstrated the cost-effectiveness of the UK standard compression bandaging system compared with a widely used alternative.
Value of information	Showed that commissioning of future RCTs of pentoxifylline is unlikely to be an efficient use of research resources.

Table 3. Summary of contributions to knowledge

5.2 Recommendations for future research

A series of Cochrane systematic reviews in wound care management, of which VLUs are an aspect, highlighted both the poor methodological standards of existing clinical and economic evidence in this area and a need for high quality primary research in a number of elements of VLU care.[7-9] Research priorities identified include the need to:

(i) investigate the differential effectiveness of therapeutic devices for VLUs (compression therapies, wound dressings, topical agents, pharmacological treatments, debriding agents, an other treatments) licensed for the same indication;

(ii) assess the relative cost-effectiveness of alternative treatment strategies;

(iii) explore the HRQoL consequences associated with alternative treatments for VLUs;

(iv) improve methodological standards of primary research in VLUs according to international guidelines for the design, conduct and analysis of RCTs.

In order to define what constitutes the efficient management of VLUs, a single comparative analysis of the clinical and cost-effectiveness of all existing external high compression systems is required. This analysis would require an evidence-based categorisation of all external high compression systems available. Currently, in addition to the level of compression delivered, a distinction is made between systems according to the number of layers of compression required (*single*, *2-layer*, *3-layer*, *4-layer*), the degree of elasticity (elastomeric, non-elastomeric)·, and the method of

delivery (bandage, stockings, rigid cast, pneumatic). Evidence of the rationale for attributing differential effectiveness on the basis of these characteristics is needed.

Recently, Bayesian statistical methods have been proposed as a suitable approach for CEA.[95] There are a number of reasons for this: i) decision analysis is intrinsically Bayesian; ii) probability statements relevant for decision-making "*the probability that an intervention is clinically and / or cost-effective given the current level of evidence…*" are only valid in a Bayesian framework; iii) Bayesian methods consent the analysis of newly acquired evidence in the light of previously existing data; iv) with the development of powerful numerical methods, i.e. Markov chain Monte Carlo simulations, previously unsolved analytical problems can be resolved within a Bayesian paradigm.[127]

The potential contribution of these methods to the area of VLUs is considerable. For example, in the absence of head to head comparisons, Bayesian mixed treatment comparisons (of which indirect comparisons are a specific case), can be used to allow the synthesis of evidence from available studies in which the treatments of interest have been directly compared with some other treatments, in paired or multiple comparisons. [110] For instance, if the comparison between *short-stretch* bandage and *2-layer* bandage is of interest, data on comparisons between *2-layer* vs. *4-layer*, *short-stretch* vs. *2-layer* vs. *4-layer*, and *short stretch* vs. *4-layer* could be used to derive an estimate of the relative effectiveness of *short stretch* vs. *2-layer* bandages.

Similarly the relative clinical and cost-effectiveness of dressings, debriding agents, topical agents, pharmaceutical therapies and other treatments available for VLUs requires further investigation. In this case an evidence-based reclassification of available treatments will also be required.

Finally, there is a need to ascertain the contribution of concomitant diseases in the HRQoL of people with VLUs. This issue needs to be addressed in future studies in which the presence or development of concomitant chronic conditions is carefully documented.

Appendix A

Original research

Does length of questionnaire matter? A randomised trial of response rates to a mailed questionnaire

Cynthia Iglesias, David Torgerson[1]

Department of Health Studies and [1]Centre for Health Economics, University of York, UK

Objective: To assess whether length of questionnaire affects response rates.

Methods: A quasi-randomised trial of women aged 70 years and over in a general practice in England. Three questionnaires of different lengths: a clinical questionnaire (four pages); the same questionnaire plus the EuroQol (five pages); the same questionnaire plus the SF-12 (seven pages). The impact of length on the proportion of returned questionnaires and item completion rates was assessed.

Results: In total, 847 questionnaires were mailed; response rates were 49%, 49% and 40% to the short, medium and long questionnaires, respectively. This difference was statistically significant when the short questionnaire was compared against the longest instrument (9% difference; 95% confidence interval (CI) of difference=0.3% to 16.6%). Item completion rates for the clinical questionnaire did not differ. Respondents did not differ in age or self-reported health status between the three groups.

Conclusions: Increasing the length of a questionnaire from five to seven pages reduces response rates from women aged 70 years and over. However, lengthening a questionnaire does not seem to affect the quality of responses to questions near the front of the questionnaire.

Journal of Health Services Research & Policy Vol 5 No 4, 2000: 219–221

Introduction

Many researchers are resistant to increasing the length of questionnaires because of the perception that they may lead to poorer response rates. This perception has been supported recently when a randomised trial comparing response rates to a clinical questionnaire plus the EuroQol with a clinical questionnaire plus the SF-36 showed a modest (5%) but statistically significant reduction in questionnaire response.[1] Although it is commonly thought that questionnaire length influences response rates, there is limited evidence to support this view. A review of patient satisfaction surveys found response rates were unaffected by questionnaire length.[2] Another review[3] found only two trials of questionnaire length in the context of health care research, both of which found no effect of questionnaire length.[4] However, non-British and non-health studies have found evidence supporting a negative relationship between questionnaire length and response rates.[6,7] Given the importance of the topic, surprisingly few studies have explored the effect that length of questionnaire has on response rate.

Cynthia Iglesias MSc, Research Fellow, Department of Health Studies, **David Torgerson PhD**, Senior Research Fellow, Centre for Health Economics and Department of Health Studies, University of York, York YO10 5DD, UK

Correspondence to: CI.

Researchers from different backgrounds, such as health economics and clinicians, often require different quality of life instruments. For example, economists like to have a measure of utility – such as the EuroQol – whereas other health services researchers may prefer instruments such as the SF-36. However, inclusion of a lot of instruments increases questionnaire length. Therefore, we would like to know whether inclusion of multiple quality of life instruments on top of a clinical questionnaire is detrimental to response rates.

Methods

In a questionnaire survey of osteoporotic risk factors among women aged 70 years or over in Yorkshire, women were asked about fracture risk factors and willingness to participate in a randomised trial of hip protectors. Women were encouraged to return the questionnaires, irrespective of whether they wanted to take part in the trial. General practitioners provided names and addresses of those patients who they thought could answer the questionnaires or who were not bed- or chair-bound.

Four different versions of the questionnaire were designed. The first consisted of clinical questions relating to osteoporosis risk factors, the first question of the SF-12 and the patient's willingness to participate into the study (four A4 pages). The second included the EuroQol (without the visual analogue scale) in addition to the clinical questions (five A4 pages). The third

questionnaire comprised the clinical questions, the EuroQol and the original version of the SF-12. Finally, the fourth questionnaire was the same as the third but used the newest version of the SF-12 (seven A4 pages).

To produce equivalent groups, women were randomised by alternation, which, if strictly adhered to, results in comparable groups.[8] The questionnaires were put into an opaque brown envelope, were alternated, and given to a member of the practice staff to be addressed and mailed. All the patients received a package containing one of the questionnaires, an information letter describing the aim of the study, and a prepaid return envelope. No reminders were sent. Completed questionnaires were returned to the Centre for Health Economics.

To detect a difference of 15% in response rate (50–65%) at 80% power ($2P$=0.05) for a two-group comparison requires a total of approximately 366 people (i.e. 183 in each group). Results were considered statistically significant with a P value of 0.05 or less. Each patient received a questionnaire between February and April 1999.

Results

A total of 856 questionnaires were sent out: 278, 278, 150 and 150 of types 1, 2, 3 and 4, respectively. Initial sample size was reduced by nine patients who were ineligible for the trial (death, dementia, moved away from the study area). In the analysis a final sample size of 847 subjects was considered.

The mean age for the study population was 77.7 years (standard deviation (SD)=6.8). As the two versions of the SF-12 had only trivial differences (i.e. change from vertical to horizontal format in one question and addition of two response categories for two questions), which did not affect its length, data on questionnaires 3 and 4 were pooled. Therefore, the questionnaires were classified into three categories: short, medium and long. A comparison of the three groups with respect to age and general health status revealed no obvious differences between the groups (Table 1). Similarly, a comparison between the medium and long questionnaires in terms of EuroQol scores (0.87) showed no difference

Table 2 shows that the longest questionnaire had a response rate about 9% lower than those of the shorter questionnaires. We found no significant relationship between length of questionnaire and the number of missing items in the clinical questionnaire (Table 3). A comparison between the medium and long questionnaires in the number of completed EuroQol instruments showed a slight increase in fully completed EuroQols for the longer questionnaire: 89.6% (n=121/135) and 93.3% (n=111/119) for medium and long questionnaires, respectively (95% CI of difference: −10.5% to 3.19%). Finally, questionnaire length had no effect on the numbers of patients expressing a willingness to participate in the trial: 7.2% (20/276), 9.0% (25/277) and 8.8% (26/294) for short, medium and long questionnaires, respectively.

Discussion

This is the first reported trial, within a UK population, which shows that increasing the length of a questionnaire lowers the response. Our overall response rate was 46%, comparatively smaller than in similar studies (62.1–79%).[4,5,10] We were unable to use reminders and this will partially explain our low response. Furthermore, an aim of our study was to recruit patients into a randomised trial, and those women who would have normally responded to a survey may not have responded if they did not wish to be included.

Lengthening the questionnaire did not affect the number of missing clinical items. This might not have happened had the clinical items been placed at the back. Although the response rate to the long questionnaire was somewhat lower than the medium questionnaire, the completion rate for the EuroQol component was similar. Thus, if a respondent was prepared to return a long questionnaire she was likely to complete the EuroQol instrument in the middle of the questionnaire. This supports the findings of a similar study undertaken to evaluate various measures likely to increase response rates.[11] Of importance to health economists is that lengthening a clinical questionnaire by the trivial amount required to include the EuroQol has no discernible effect on response rates.

Although our results suggest that a reduction in response rates is due to increased questionnaire length, they need to be placed in the context of our study population. Our results may only be applicable to older women. Furthermore, questionnaire length may not

Table 1 Analysis of variation among different types of questionnaires with respect to age and health status

Questionnaire type	Age (years)		Health status	
	Mean (SD)	Number[a]	Proportion of good, very good or excellent (%)	Number[a]
Short	77.9 (7.8)	122	33.7	133
Medium	77.6 (6.5)	117	34.4	134
Long	77.6 (6.1)	102	32.0	119
Total	77.7 (6.8)	341	76.2	386
	t=0.24	Significance 0.691	χ^2=0.79	Significance 0.674

[a]Variations due to missing items. SD, standard deviation.

Table 2 Interactions between retrieval rate and questionnaire length

Response rates (%)				*P* value	95% CI of difference[a]
Short	Medium	Long	Overall		
48.9 (135/276)	48.7 (135/277)	40.5 (119/294)	46 (389/847)	0.04	0.3 to 16.6
$\chi^2 = 4.1017$					
Odds ratio[b]=1.19		CI=(1.005 to 1.401)			
Odds ratio[c]=1.12		CI=(1.018 to 1.239)			

[a]Comparison between the response rates to short and long questionnaires.
[b]Odds ratio of long questionnaire compared with short questionnaire retrieval rates.
[c]Odds ratio of long questionnaire versus short+medium questionnaire retrieval rates.
CI, confidence interval.

Table 3 Partial non-response to clinical questions across questionnaires: missing items to clinical questions

Partial response	Type of questionnaire		
	Short	Medium	Long
No	83 (61.5%)	83 (61.5%)	71 (59.7%)
Yes	52 (38.5%)	52 (38.5%)	48 (40.3%)
Total	135	135	119
χ^2=0.11462	Significance=0.9443		

have a linear relationship with response rate: there could be a threshold effect in which response rates suddenly drop off. Given the paucity of research into the effect of questionnaire length on response rates more studies are required in a range of different populations.

Acknowledgements

We thank the two anonymous referees for their helpful comments.

References

1. Dorman PJ, Slattery J, Farrel B, Dennis M, Sandercock P. A randomised comparison of the EuroQol and Short Form-36 after stroke. BMJ 1997; 315: 461–463
2. Sitzia J, Wood N. Response rate in patient satisfaction research: an analysis of 210 published studies. International Journal of Quality of Health Care 1998; 10: 311–317
3. McColl E, Jacoby A, Thomas L, Soutter J, Bamford C, Thomas R et al. Designing and using patient and staff questionnaires. In: Black N, Brazier J, Fitzpatrick R, Reeves B, eds. Health services research methods: a guide to best practice. London: BMJ Books, 1998
4. Jacoby A. Possible factors affecting response to postal questionnaires: findings from a study of general practitioner services. Journal of Public Health Medicine 1990; 12: 131–135
5. Cartwright A. Some experiments with factors that might affect the response of mothers to a postal questionnaire. Statistics in Medicine 1986; 5: 607–617
6. Hansen R, Robinson L. Testing the effectiveness of alternative foot-in-the-door manipulations. Journal of Marketing Research 1980; 17: 359–364
7. Powers DE, Alderman DL. Feedback as an incentive for responding to a mail questionnaire. Research in Higher Education 1982; 17: 207–211
8. Roszkowski MJ, Bean AG. Believe it or not – longer questionnaires have lower response rates. Journal of Business and Psychology 1990; 4: 495–509
9. Chalmers I. Assembling comparison groups to assess the effects of health care. Journal of the Royal Society of Medicine 1997; 907: 379–386
10. Lund E, Gram I. Response rate according to title and length of questionnaire. Scandinavian Journal of Social Medicine 1998; 26: 154–160
11. Eaker S, Bernström R, Bernström A, Adami H, Nyren O. Response rate to mailed epidemiologic questionnaires: a population-based randomised trial of variations in design and mailing routines. American Journal of Epidemiology 1998; 147: 74–82

Improving the measurement of quality of life in older people: the York SF-12

C.P. IGLESIAS, Y.F. BIRKS and D.J. TORGERSON

From the Department of Health Studies, University of York, York, UK

Received 3 August 2001 and in revised form 9 October 2001

Summary

To assess whether changing the layout of the SF-12 affected item response rates, we tested two SF-12 formats in a quasi-randomized trial of women aged ⩾70 years in two general practices in North Yorkshire. The modified version of the SF-12 ('York SF-12') converted the 'stem and leaf' format of some questions to individual items. We assessed the effect of the two types of questionnaires on item response rates. The difference in overall response rates to the two questionnaires (York SF-12 26.8%; SF-12 29.5%) was not statistically significant (95%CI −1.88% to 7.22%). However, the modified SF-12 had a statistically significantly lower item non-response rate of 8.5%, compared with the 26.6% of the SF-12 (95%CI 11.1%–25.1%). Chronbach's alpha reliability scores for the York SF-12 were also slightly better than for the older version. The York version of the SF-12 is an improvement on the original questionnaire. We recommend that the York SF-12 be used in preference to the SF-12 when surveying an older population.

Introduction

The SF-36 is a widely used quality-of-life instrument. However, its length could affect response rates, particularly in older people. For instance, we previously found an 8% reduction in response rates in a population survey when a questionnaire was extended from 4 to 7 pages of A4,[1] while Dorman and colleagues using a population from a randomized trial noted a 5% reduction in response rates when the SF-36 was compared against the much shorter EuroQol questionnaire.[2] Because of possible problems with response rates and data entry burden a shortened version of the SF-36 (the SF-12[3]) is popular. Recently, both of these instruments have been amended. For instance, some questions that had a binary response have been replaced with a Likert response scale. In addition, the layout for some questions has been changed from a vertical format to a horizontal layout.

We are using the latest version of the SF-12 in a number of ongoing randomized trials. However, the SF-12 (and the SF-36) has a confusing 'stem and leaf' layout of some of its items. Some questions are preceded with a general phrase such as 'How much during the last month:', which is then followed by up to three specific questions (Figure 1). We noticed in a pilot study that this layout is confusing to older respondents. In the context of a pilot study for a large randomized control trial of patients with chronic leg ulcers, eleven elderly patients were asked to complete four different quality-of-life questionnaires, presented in the following order, Euroqol, SF-12, pain, and ulcer-specific questionnaires.

Even though patients were asked to respond the questionnaires on their own, they often asked for assistance with questions in the SF-12. They had difficulties working out the way they should respond to the 'stem and leaf' questions. When it was explained to patients that the researcher could not help, since this could bias their responses, patients tended to either miss items (questions) from

Address correspondence to Ms C.P. Iglesias, Department of Health Studies, University of York, York YO10 5DQ. e-mail: cpiu1@york.ac.uk

a

These questions are about how you feel and how things have been with you during the **past 4 weeks**. For each question, please give the one answer that comes closest to the way you have been feeling. How much of the time during the **past 4 weeks**...

	All of the time	Most of the time	Some of the time	A little of the time	None of the time
d) have you felt calm and peaceful?	☐	☐	☐	☐	☐
e) did you have a lot of energy?	☐	☐	☐	☐	☐
f) have you felt downhearted and low?	☐	☐	☐	☐	☐

b

1. How much during **the last month** have you felt calm and peaceful?

*(Please circle **one number only**)*

All of the time 1 | Most of the time 2 | Some of the time 3 | A little of the time 4 | None of the time 5

2. How much during **the last month** did you have a lot of energy?

*(Please circle **one number only**)*

All of the time 1 | Most of the time 2 | Some of the time 3 | A little of the time 4 | None of the time 5

3. How much during **the last month** have you felt downhearted and low?

*(Please circle **one number only**)*

All of the time 1 | Most of the time 2 | Some of the time 3 | A little of the time 4 | None of the time 5

Figure 1. Example of **a** the current 'stem and leaf' layout and **b** the alternative layout proposed for the SF-12.

the questionnaire, tick the same item twice, or tick the descriptions of the response categories rather than the appropriate box (Iglesias, unpublished data). This results in loss of data and also leads to data entry problems. When a respondent ticks two categories, not only is it impossible to ascertain which is the true response, but when data-entry is through a scanner an error message is produced, requiring the questionnaire to be manually checked, which increases the cost and complexity of data entry. Similarly, if the respondent ticks the description of the response category rather than the box, this again leads to problems with automated data entry. Because of these problems, we decided to amend the layout of the SF-12 without altering the existing questionnaire's length (2-sides of an A4 page) and undertake a randomized comparison of our version on its response rates and validity.

Methods

We amended the SF-12 questionnaire by converting [illegible] individual items. This version was piloted among a small convenience sample of patients at a leg ulcer clinic. After some slight revisions we decided to test the 'York SF-12' in a randomized trial.

We undertook an opportunistic study in the context of recruiting women aged 70 years and over for a randomized trial of hip protectors for fracture prevention. We mailed out 1500 questionnaires, 750 of each version, York SF-12, and SF-12. This was to detect at least an 8% difference in response rate at 80% power ($2p \leqslant 0.05$). The EuroQol without the analytic scale was also included in both questionnaires. Questionnaires were placed in an opaque sealed envelope with a prepaid reply envelope. To produce equivalent groups, we used alternation, which if strictly adhered to, results in comparable groups.[4] The sealed envelopes were then sent to participating general practices. Staff at the general practices put address labels on the envelopes and posted them to the participants. No reminders were or could be sent. Completed questionnaires were returned to the Department of Health Studies.

Statistics

Comparisons in return and item completion rates, and general health status of respondents between the two types of questionnaires were performed using χ^2 tests. We also compared item-response rates between the EuroQol and the two versions of the SF-12 within questionnaires using the McNemar test for paired data. To assess the validity of the new version, we explored any possible psychometric differences between the York SF-12 and the SF-12 questionnaires by principle components analysis, using a varimax orthogonal rotation with an exclusion criteria of 0.5. The adequacy of the data for factor analysis was identified using Kaiser Meyer Olkin (KMO) test and Bartlett's test of sphericity.[5] Chronbach's alpha[6] was used to explore questionnaires' internal reliability.

Results

Out of the 1500 questionnaires mailed out, 422 questionnaires were returned, 221 (29.5%) and 201 (26.8%) for the old and new versions, respectively. No statistically significant difference ($p = 0.779$) was found in a comparison of the general health status between the respondents to either version of the SF-12. Differences in overall response, missing item, and single item response rates per questionnaire are described on Table 1. No statistically significant difference in the overall response rates were identified.

As Table 1 shows, there were no statistically significant differences in the number of returned questionnaires (i.e. confidence interval of mean difference includes zero). Conversely, the number of missing items in either version of the questionnaire was statistically significant for the majority of questions with the 'stem and leaf' layout, as well as overall non-response to any item.

We also compared the item response rates between the two versions of the SF-12 with the EuroQol. As expected, there were no statistically significant differences in item non-response rates for the EuroQol when combined with either version of the SF-12, 3.1% and 2% for SF-12 and York SF-12 respectively, (95%CI −1.85% to 4.14%). However, the difference (6.5%) in the modified SF-12 item non-response compared with the EuroQol, although statistically different, (95%CI 2% to 8%, $p < 0.004$), was substantially lower than the 23.8% difference between the old SF-12 and the EuroQol (95%CI 17.8% to 29.8%, $p < 0.001$).

The KMO and Bartlett's tests were both satisfactory for both versions of the questionnaire (0.93049 and 2060.2787, $p < 0.0001$ for the York SF-12; 0.90573 and 1455.5486 $p < 0.0001$ for the old SF-12). Subjecting both questionnaires to factor analysis yielded the expected two-factor solution. The internal reliability of the York SF-12 was 0.94 and 0.91 for the physical health and mental health domains, respectively, measured using Chronbach's alpha. This was slightly better than the standard SF-12, which gave reliability estimates of 0.90 and 0.88 for the same two factors.

Discussion

The SF-12 is an increasingly popular quality-of-life questionnaire. However, we found that its layout caused confusion among the older population in our studies. It would appear that this may be due to

Table 1 Differences in response rates between the standard SF-12 and the York SF-12

	Standard SF-12, *n* (%)	York SF-12, *n* (%)	Difference (95%CI)
Overall response rate	221 (29.5%)	201 (26.8%)	2.7% (−1.88% to 7.22%)
Item response rates			
Q1	219 (99.1%)	199 (99%)	−0.1% (−1.95% to 1.77%)
Q2	**214 (96.8%)**	**199 (99%)**	**2.2% (−0.5% to 4.86%)**
Q3	**206 (93.2%)**	**199 (99%)**	**5.8% (2.20% to 9.38%)**
Q4	**209 (94.6%)**	**199 (99%)**	**4.4% (1.15% to 7.72%)**
Q5	**191 (86.4%)**	**198 (98.5%)**	**12.1% (7.27% to 16.9%)**
Q6	**209 (94.6%)**	**195 (97.5%)**	**2.9% (−1.36% to 6.25%)**
Q7	**198 (89.6%)**	**196 (97.5%)**	**7.9% (3.35% to 12.5%)**
Q8	209 (94.6%)	198 (98.5%)	3.9% (0.51% to 7.36%)
Q9	**202 (91.8%)**	**198 (98.5%)**	**6.7% (3.05% to 11.2%)**
Q10	**203 (91.9%)**	**197 (98%)**	**6.1% (2.06% to 1.02%)**
Q11	**205 (92.3%)**	**198 (98.5%)**	**6.2% (1.94% to 9.55%)**
Q12	214 (96.8%)	194 (96.5%)	−0.3% (−3.74% to 3.11%)
Non-response to any item	58 (26.6%)	17 (8.5%)	18.1% (11.1% to 25.1%)

Items in bold are questions in the 'stem and leaf' format in the standard SF-12.

the 'stem and leaf' format of some of the SF-12 items. Converting these items into individual items led to a statistically significant decrease in the number of missing items. The new layout of the York SF-12 appears to have an unambiguous factor structure with good reliability. Importantly, it improves item response by clarifying the format, resulting in fewer missing values, while showing a similar reliability.

We accept that our overall response rates for both studies were relatively low. This is probably explained by the fact that the main aim of our survey was to recruit high-risk women to a randomized trial of hip protectors. However, there was only a small difference in overall response rates between the two questionnaires, which was not statistically significant. Further, if our study group was relatively highly motivated to complete the questionnaire, this would suggest that the York SF-12 could perform even better, compared with the standard SF-12, in a less-motivated population. It is interesting to note that in a trial comparing the standard version of the SF-36 with the EuroQol, 28% of the returned SF-36 responses had missing data, which is similar to the 27% we found using the standard SF-12.[2] It may be helpful to modify the SF-36 in a similar manner.

In conclusion we have modified the SF-12 and undertaken a trial to test its item response rate and its reliability. We are currently exploring the equivalence of the Physical Component Score and Mental Component Score between the two versions of the SF-12.[7] However on the basis of the present analysis, we would recommend using the York modified SF-12, particularly among older people.

Acknowledgements

This research was undertaken as part of a project funded by Procter & Gamble Pharmaceuticals and Aventis Pharma.

References

1. Iglesias C, Torgerson D. Does length of questionnaire matter? A randomized trial of response rates to a mailed questionnaire. *J Hlth Serv Res Policy* 2000; **5**:219–21.
2. Dorman PJ, Slattery J, Farrell B, Dennis MS, Sandercock PAG. A randomised comparison of the EuroQol and Short Form-36 after stroke. *Br Med J* 1997; **315**:461–63.
3. Jenkinson C, Layte R, Jenkinson D, Lawrence K, Petersen S, Paice C, Stradling J. A shorter form health survey: can the SF-12 replicate results from the SF-36 in longitudinal studies? *J Pub Hlth Med* 1997; **19**:179–86.
4. Chalmers I. Assembling comparison groups to assess the effects of health care. *J Roy Soc Med* 1997; **907**:379–86.
5. Norussi MJ. *SPSS-X advanced Statistics Guide*. New York, McGraw Hill, 1985.
6. Bland JM, Altman DG. Statistics notes: Chronbach's alpha. *Br Med J* 1997; **314**:572.
7. Ware JE, Kosinski M, Keller SD. *How to score the SF-12 Physical & Mental Health Summary Scales*, 3rd edn. Lincoln RI, QualityMetric Inc, 1998.

Quality of Life Research (2005) 14: 1705–1718
DOI 10.1007/s11136-005-2751-9

Quality of life of people with venous leg ulcers: A comparison of the discriminative and responsive characteristics of two generic and a disease specific instruments

Cynthia P. Iglesias, Yvonne Birks, Elizabeth A. Nelson, Elizabeth Scanlon & Nicky A. Cullum
Department of Health Sciences, University of York, North Yorkshire, UK (E-mail: cpiu1@york.ac.uk)

Accepted in revised form 18 February 2005

Abstract

Background: Venous leg ulcers are an important source of morbidity in society. Measuring the impact of leg ulcers on quality of life is important within clinical and economic evaluations. In this study we report a validation study of the leg ulcer disease specific Hyland questionnaire and compare its discriminative and responsive characteristics to general health quality of life measures: the SF-12 and EQ-5D. *Methods*: HRQoL of venous leg ulcer patients from 9 UK regions was measured using SF-12, EQ-5D and Hyland, at baseline and every three months for 1 year. Psychometric analysis was used to confirm the validity of the Hyland questionnaire. Quarterly scores for all instruments were calculated. Effect size and standardised mean difference were used to investigate the responsiveness to ulcer healing and discriminative abilities of the instruments. *Results*: Three hundred and eighty seven individuals were recruited into the VenUS I study. Baseline health related quality of life data from the study participants suggested a two factor solution for the Hyland. This questionnaire was associated with small and moderate ability to discriminate individuals according to age, mobility, initial ulcer size and ulcer duration. SF-12 and EQ-5D had good evaluative properties; both instruments were responsive to changes in HRQoL after ulcer healing. High levels of bodily pain were reported in the SF-12 questionnaire, whilst only minor ulcer related discomfort was reported in the Hyland. *Discussion*: SF-12 and EQ-5D are suitable for exploring dimensions of health related quality of life in people with chronic venous leg ulceration. The responsiveness to healing of the Hyland questionnaire is unclear. We would recommend the use of generic instruments for the measurement of HRQoL in patients with venous leg ulcers.

Key words: EQ-5D, Health related quality of life, Hyland, SF-12, Venous leg ulcers

Abbreviations: HRQoL – health related quality of life; SF-12 – Short Form 12; EQ-5D – Euroqol 5 dimensions; ES – effect size; MSD – mean standard difference; PCS – Physical Component Score; MCS – Mental Component Score

Introduction

Health related quality of life (HRQoL) in people with leg ulceration has been investigated in a number of ways including indepth qualitative studies; cross-sectionally using generic HRQoL instruments (e.g. Short-Form 36 (SF-36), Nottingham Health Profile (NHP)) and/or leg ulcer specific measures in people with leg ulcers and/or people of a similar age without leg ulcers; and prospectively within evaluations of therapies [1–9].

Several studies have used inductive approaches such as phenomenology to inquire about life with a leg ulcer from the patient perspective [1–3]. Living with chronic pain is a theme that consistently emerges from these studies and yet venous

leg ulcers were traditionally not regarded as painful [2–4]. Other themes which recur are the restrictions to social, leisure and work activities [1–3], the hope and despair experienced throughout the long healing trajectory [1, 3], and the restrictions ulceration places on clothing and footwear that can be worn [3, 4].

In the context of the VenUS I study (a large UK multicentre pragmatic randomised controlled trial of two bandages in the treatment of chronic venous leg ulcers), the health related quality of life of the participants was investigated using two generic and one disease specific quality of life questionnaires. At the commencement of the study, three disease specific measures of HRQoL for venous ulcers were available: the Hyland [10], the Chronic Venous Insufficiency Questionnaire (CIVIQ) [11], and the Freiburger Questionnaire of QoL in Venous Diseases (FLQA) [12].

Whilst CIVIQ and FLQA are both health measures for venous insufficiency (a spectrum of disorders ranging from varicose veins through to extensive ulceration of the leg), with good evidence of their validity and reliability [11, 12], Hyland was claimed to measure HRQoL of people with ulceration [10]. The Hyland scale had been constructed using sound psychometric methodology, however given the number of items in the scale, the sample size for the factor analysis in the original study was small [10]. The VenUS I trial presented an opportunity to confirm the factorial structure, validity and reliability of the Hyland questionnaire in a larger sample of UK individuals.

Of the most widely used generic quality of life instruments, the SF-12 and EQ-5D were selected because of their conciseness and relative simplicity. The SF-12 measures physical and mental components of health related quality of life [13]. The EQ-5D was constructed to measure levels of utility associated with specific health status [14].

This paper presents the results of the validation of the Hyland questionnaire as well as a comparison of the three instruments in terms of their ability to discriminate between individuals who are experiencing different levels of health related quality of life (discriminative properties), and their ability to detect changes in HRQoL that occur as a result of compression therapy, i.e. the presence/absence of an open ulcer (evaluative properties/responsiveness).

Methods

Individuals were asked to complete a survey describing their HRQoL during the three months preceding recruitment into the VenUS I trial (the baseline assessment), and at quarterly intervals thereafter. The survey was designed to be completed by study participants and returned to the trial office using a pre-paid envelope. Individuals were asked to complete up to three different HRQoL instruments: EQ-5D, SF-12 and Hyland. The SF-12 and EQ-5D had to be completed by all participants irrespective of the condition of their leg ulcer. In a pilot study of the three HRQoL instruments to be used in the VenUS I trial, the relevance of a number of questions in the Hyland questionnaire after ulcer healing was questioned by the participants. As a result, in order to prevent confusion in completing this questionnaire, the research team decided to administer the Hyland only to participants with an open venous leg ulcer.

Hyland questionnaire

The Hyland questionnaire comprises 34 questions in three sections. In the first section, individuals are asked whether they have been hospitalised or housebound because of their leg ulcer, and to describe the current progress of their ulcer on a visual scale. The second section comprises four items: leg ulcer pain, sleep discomfort because of leg ulcer, time thinking about the ulcer, and time spent helping the ulcer heal. The third section comprises a list of 29 HRQoL items concerning functional limitation, dystrophic mood and treatment associated with the presence of an open leg ulcer, e.g. *'I worry about shopping trolleys'* [10]. In a pilot study of the three quality of life instruments to be used in this study, individuals with a venous leg ulcer identified a number of potential limitations with the original version of the third section of the Hyland questionnaire. Subsequently questions in this section were rephrased using a more colloquial language. Furthermore, the four-category 'never' to 'always' scale originally used in the third section of Hyland was replaced by a simple yes/no scale. While reduced response scales can be associated with poorer longitudinal responsiveness it was considered that a simpler

response scale would minimise single item non-response [15].

Given the small sample size with which the original one factor solution for the third section of the Hyland questionnaire was determined (50 individuals), baseline data from this section were used to conduct an exploratory factor analysis using principal-axis factoring in SPSS 10 [16]. The scree plot provides a more conservative estimate of the number of factors and produces less shared variance, so this criterion was used as an indication of the number of factors both here and in the original scale construction, as opposed to the criterion of those with an Eigenvalue > 1 [17]. Principal axis factoring using a Varimax orthogonal rotation was used. This form of factor analysis seeks to establish the least number of factors to account for the common variance in a set of variables, and is more conservative than the principal components analysis used in the original scale construction where evidence of a one factor solution was identified [18]. An exclusion criterion of 0.3 was employed [19]. Cronbach's α was used to explore the internal reliability of the Hyland tool [20].

A simple scoring system of adding one point for every 'yes' response to a negative attribute was used to score the third section of the Hyland. Consequently, higher scores indicate poorer HRQoL the worst possible scores for the physical and mental factors are 12 and 14, respectively.

EQ-5D questionnaire

The EQ-5D is a generic measure of health status, where health is characterised on five dimensions (mobility, self care, ability to undertake usual activities, pain, anxiety/depression) and it is useful within economic evaluations because it has ratio properties [14]. Individuals are asked to describe their health on each dimension using one of three levels: no problems, moderate problems and severe problems. Each response locates a person into one of 243 mutually exclusive health states, each of which has previously been valued on the 0 (equivalent to dead) to 1 (equivalent to good health) scale based on interviews with a sample of 3395 members of the UK public [21]. The EQ-5D 'utility' scores range from −0.57 to 1. Utility scores for each patient were calculated using the responses to the EQ-5D at baseline and every three months afterwards during the first year of follow up.

SF-12 questionnaire

In a pilot study of the VenUS I trial HRQoL instruments, it was noticed that the original 'stem and leaf' layout of the SF-12 presented some difficulties to older respondents. The layout of the SF-12 was then modified and tested in a population of women aged 70-year and over. The new layout was associated with lower item non-response and a reliability score slightly better than that of the older version [22]. This new layout was used throughout the VenUS I follow up period.

The scores for the physical and mental components of the SF-12, as well as the scores for its 8 individual health dimensions; general health, physical functioning, role physical, role emotional, bodily pain, mental health, vitality, and social functioning, were calculated at baseline and every three months for the first year of follow up [13]. To facilitate the interpretation and comparison of our results with those from other studies, all scores were norm-based to the average mean values of the general US population, i.e. scores were transformed to be distributed with a mean of 50 and a standard deviation of 10. Any values below 50 indicate that the physical or mental components of the population under study are below the average of the general US population; the opposite is true for values above 50.

Discriminative and evaluative properties (responsiveness to ulcer healing)

The ability of the three HRQoL instruments to discriminate individuals according to age (age $\leqslant$ 71 years/age > 71 years), degree of mobility (fully/partially mobile), initial ulcer size (size $\leqslant$ 4 cm^2/size > 4 cm^2), and duration of current ulcer (duration $\leqslant$ 3 months/duration > 3 months), at baseline was explored using the effect size (ES). For continuous variables, i.e. age, ulcer size and ulcer duration, the cut-off point to split the study sample into two groups was determined by the sample's mean or median in the case of highly skewed variables, such as ulcer's size and duration. ES was calculated as the ratio of the difference in means between the two groups, de-

fined by the characteristic of interest (i.e. age, mobility, etc), to the full sample standard deviation at baseline [23].

The evaluative properties of the Hyland questionnaire with respect to venous leg ulcer healing could not be explored. The relevance of many of this instrument's items after ulcer healing was unclear. Thus, this questionnaire was only administered to individuals with an open ulcer. The responsiveness of the two generic instruments to changes in HRQoL associated with the healed/not healed state with respect to baseline scores was investigated using the ES and standardised response mean (SRM). In this case, ES was estimated as the ratio of the mean change between baseline and three-month scores to the standard deviation of the scores at baseline [23]. Similarly SRM is the ratio of the mean change between baseline and 3-month scores to the standard deviation of the difference in means [23].

Repeated measures model

Patterns of missing data were explored following the procedure described by Fairclough [24]. Logistic regression was used to investigate the association of missing data with baseline covariates and previously observed HRQoL. This analysis indicated that the only statistically significant predictors of non-response at any specific follow up time were whether they had previously responded, thus suggesting that the pattern of missing data was not missing completely at random but missing at random [24]. Descriptive analysis also indicated that the patterns of missing data were non-monotonic. Consequently, a Bayesian two level random effect model was constructed in WinBugs1.4 software package to estimate mean scores for SF-12, EQ-5D and Hyland at each follow up point, as well as to perform hypothesis testing [25–27]. The first level of the model was given by the quality of life scores reported at different follow up points by the same individual, these in turn were clustered within individuals which defined the second level of the model. Vague priors were fitted to all unknown parameter within the model [28]. Missing data were treated as parameters thus imputed/predicted with the posterior predictive distribution of the mean scores at each time point [29].

An exploratory analysis indicated no statistically significant differences in HRQoL scores between bandage groups over time for any of the three different instruments. In order to simplify the presentation of the analysis the data from both bandage groups were combined in a single dataset of individuals with venous leg ulcers.

Results

Three hundred and eighty seven individuals were recruited in the VenUS 1 study. Demographic and ulcer specific characteristics of the participants at baseline are provided in Table 1. The clinical and economic analyses of the trial have been reported in detail elsewhere [30].

Response rates

The number of participants expected to have completed the Hyland, SF-12 and EQ-5D questionnaires over the first year of follow up are described in Table 2. Complete (i.e. every single item) and partial (i.e. at least one item) response rates for the three HRQoL questionnaires indicate that the response rates to SF-12 were the highest and ranged from 93% at baseline down to 61% at 12 months. Conversely, the poorest response rates were associated with the Hyland questionnaire which ranged from 92% at baseline to 49% after 1 year. Despite the EQ-5D's brevity and its position in the booklet (beginning of the HRQoL survey), its response rates were not better than those for the SF-12; range 91–59%. When only questionnaires with complete responses were considered, response rates were reduced by between 1% and 24%.

Factor analysis

Since the original analysis had described the third section of the Hyland as unidimensional scale, a one factor solution was examined; all 28 items loaded. Items 14 (*'Because of my ulcer I try to keep away from cats etc'*) and 19 (i.e. *'I think my ulcer will not get better'*) loaded below a 0.3 criterion. The factor explained 28.75% of the variance and had an internal reliability of 0.88.

While six factors had Eigenvalues above 1.0, examination of the scree plot indicated a two-factor

Table 1. Baseline demographic variables and leg ulcer characteristics

Continuous variables	N	Mean (SD) [range]
Patient characteristics		
Age (years)	387	71.6 (13.21) [23–97]
Height (m)	379	1.7 (0.11) [1.4–2.0]
Weight (kg)	377	79.8(19.8) [33.1–142.4]
History of ulceration		
Duration of ulceration (years since onset)	372	3* [0–75]
Episodes (since onset)	375	2* [0–64]
Leg and ulcer characteristics		
Ankle circumference (cm)	380	23.9 (2.90) [16.0–34.0]
Duration of current ulcer (months)	377	3* [0–768]
Area (cm^2)	386	3.8* [0.2–254.6]
Categorical variables	**N**	**%**
Patient characteristics		
Male	159	(41%)
Female	228	(59%)
Fully mobile	238	(62%)
Walks with assistance	142	(37%)
Immobile	3	(0.77%)
Ankle mobility (full motion)	259	(67%)
Ankle mobility (reduced motion)	117	(30%)
Ankle mobility (fixed)	5	(1.3%)

*For highly skewed distributions medians are shown.

Table 2. Quarterly response rates for SF-12, EQ-5D and Hyland

	Baseline	1st quarter	2nd quarter	3rd quarter	4th quarter
Numbers eligible[a]	387	382	380	376	370
Death	0	5	7	11	17
SF-12					
Partial response	93%	77%	67%	63%	61%
Complete answers	88%	70%	62%	56%	57%
EQ-5D					
Partial response	91%	73%	66%	60%	59%
Complete answers	85%	68%	63%	59%	56%
Numbers eligible[b]	387	216	149	116	97
Healed	0	166	231	260	273
Death	0	5	7	11	17
Hyland					
Partial response	92%	68%	58%	53%	49%
Complete answers	64%	44%	36%	34%	40%

[a]Number of alive individuals in the study who should have completed an EQ-5D or SF-12 questionnaire.
[b]Number of individuals with an open ulcer who should have completed a Hyland questionnaire.

solution for the third section of the Hyland. The three highest loading items on each factor are shown in Table 3. Factor one accounted for 28.94% of the variance and factor two for a further 5.21%. In total the scale accounted for 34.15% of the variance. Unlike in the one factor solution, items 14 and 19 failed to load on either factor, consequently these items were removed from

Table 3. Factor analysis of Hyland quality of life questionnaire (section 3)

Three highest loading items on each factor

Item	Description	Factor	Loading
11	My ulcer makes getting on or off a bus difficult	1	0.735
18	My ulcer restricts where I can travel to, e.g. restricting holidays or business trips	1	0.708
9	My ulcer stops me from shopping in crowded places.	1	0.678
24	My ulcer makes me ask myself, 'Why me?'	2	0.655
23	My ulcer makes me feel depressed	2	0.654
25	Because of my ulcer it feels as though my legs/feet dominate my body	2	0.630

Double loaded items

Item	Description	Loadings	
		Factor 1	Factor 2
21	Because of my ulcer I can't be bothered to do anything	0.419	0.446
20	My ulcer gets in the way of personal relationships	0.332	0.367
16	My ulcer prevents me from wearing the type of shoes I prefer	0.337	0.365

Table 4. Correlations between baseline HRQoL scores

	Hyland practical	Hyland emotional	SF-12 physical	SF-12 mental	EQ-5D
Practical (n)	1 (247)	0.67** (247)	−0.22** (242)	−0.55** (242)	−0.61** (233)
Emotional		1 (247)	0.02	−0.60** (242)	−0.47**(233)
Physical			1 (339)	−0.12* (339)	0.25** (317)
Mental				1 (339)	0.58**(317)
EQ-5D					1 (328)

**Correlation is statistically significant at 0.01 level (2-tailed).
*Correlation is statistically significant at 0.05 level (2-tailed).

further analysis. Several items double loaded on factors. In most cases these items were sufficiently different in their loadings on the two factors for it to be apparent where they belonged. In three cases however the items were very closely loaded. There would be an argument for discarding these items as the loadings on each factor were so close but for the purposes of this study it was decided to retain them as part of the emotional factor to maintain as far as possible the structure of the original scale. These three items are shown in Table 3.

Both factors proved to be reliable using coefficient alpha to establish this. Factor one and factor two had coefficients alpha of 0.82 and 0.79, respectively. Factor one appeared to measuring some of the more practical perceptions of limitations and factor two more emotional responses to the ulcer. While the reliability was high in the one factor solution, more variance could be explained using the two factor solution with the added benefit of being able to identify perception of practical limitations as well as emotional issues. Thus, the scale was scored using two factors; a practical and an emotional one.

To test the criterion validity of the factors the practical and emotional factors were correlated with the physical and mental heath components of the SF-12, see Table 4. The practical and emotional factors of the Hyland scale correlated at 0.67. These factors also correlated in the expected direction with the SF-12 subscales: the practical factor correlated −0.22 with SF-12 physical subscale and −0.55 with the SF-12 mental health subscale, the emotional factor was not significantly correlated with the physical subscale of the SF-12 but correlated at −0.60 with the SF-12 mental health subscale.

Hyland

This section refers to the two-factor solution, however the statistics of the one-factor solution are available on request from the correspond-

Table 5. Responses to Hyland (section 2)

	Ulcer pain[a]	Sleep disturbance[b]	Thinking about ulcer[c]	Help ulcer heal[d]
Baseline				
Median [range]	3 [1–7]	2 [1–5]	3 [1–10]	5 [1–8]
N	339	335	338	168
1st quarter				
Median [range]	2 [1–7]	2 [1–5]	2 [1–10]	4 [1–8]
N	139	138	137	140
2nd quarter				
Median [range]	3 [1–5]	2 [1–4]	2 [1–10]	4 [1–10]
N	82	82	82	80
3rd quarter				
Median [range]	2 [1–5]	2 [1–5]	2 [1–10]	4 [1–8]
N	56	56	56	59
4rd quarter				
Median [range]	2 [1–5]	1 [1–5]	3 [1–10]	4 [1–8]
N	47	47	46	47

[a] *Pain:* 1. Don't notice it; 2. Uncomfortable rather than painful; 3. Hurts a little; 4. Painful; 5. Very painful; 6. Excruciatingly painful; 7. More pain that I can manage.
[b] *Sleep:* 1. Doesn't disturbs me; 2. Disturbs me only when going to sleep; 3. Sometimes wakes me up; 4. Keeps me awake a lot; 5. Keeps me awake most of the night.
[c] *Thinking:* 1. I don't think about my ulcer at all; 2. Less than 15 minutes; 3. About half an hour; 4. About an hour; 5. About an hour an a half; 6. About two hours; 7. About three hours; 8. About four hours; 9. Most of the day; 10. Most of the day and night.
[d] *Help:* 1. I don't spend any time at all trying to help my ulcer heal; 2. Less than 15 minutes; 3. About half an hour; 4. About an hour; 5. About an hour an a half; 6. About two hours; 7. Three or more hours, 8. Most of the day.

ing author. At baseline 63 (18%) participants indicated *'to have ever stayed in hospital because of their leg ulcers'*. One hundred and seventeen reported to be *'largely housebound these days'*; 81 (69%) of them indicated that this was because of their leg ulcers. Responses for section two of the Hyland are summarised in Table 5. The levels of pain, sleep disturbance, and amount of time spent looking after and thinking about their ulcer reported by the individuals on the VenUS I trial remained relatively constant over time. During the first year of follow up people reported on average that their ulcer *'hurt a little'*, *'disturbed them only when they were going to sleep'*, *'spent less than 15 minutes thinking about their ulcer'*, *and 'spent about an hour trying to help their leg ulcer heal'*.

Quarterly practical and emotional scores for individuals with an open ulcer are described in Table 6. The results indicate that individuals with an open venous leg ulcer experienced a detriment in both practical and emotional dimensions of health related quality of life. Quarterly mean scores during the first year of follow up remained relatively stable around 5 for both practical and emotional factors, i.e. a detriment of 40% in the practical and 36% in the emotional dimension. Over time no statistically significant differences with respect to baseline scores were identified with either the practical or emotional factors. The poorest scores in both factors were reported after six months of follow up.

SF-12

Mean estimates over a year's follow up of the physical (PCS) and mental components (MCS) of the SF-12 irrespective of healing status are described in Table 6. These data indicate detriments in both physical and mental components of quality of life in individuals with a venous leg ulcer. The scores for both components are below the average mean value (50) of the general US population, i.e. the health-related quality of life of the individuals in the VenUS I study was poorer than that of an average member of the US population. However, comparing baseline PCS and MCS scores, with those of an average US individual aged 75 years and over (mean age of VenUS participants was 71 years), the initial large difference in the PCS, was no longer observed [31]. On average there was a two-point difference between the mean PCS and MCS of the VenUS I

Table 6. Mean scores estimates irrespective of healing status[f]

Period	Hyland[a]		SF-12[b]		EQ-5D[c]
	Practical factor	Emotional component	Physical component	Mental component	Utility scores
Baseline (N[d])	248	248	341	341	329
Mean (SE)	5.58 (0.25)	5.10 (0.25)	35.89 (0.45)	48.45 (0.64)	0.62(0.02)
1st quarter (N[d])	95	95	266	266	261
Mean (SE)	4.64 (0.32)	4.59 (0.29)	34.92 (0.46)	50.56 (0.76)	0.69 (0.02)
Difference from baseline [95% CI]	−0.94 [−1.76 to −0.13]	−0.51 [−1.29–0.19]	−0.97 [−2.22–0.31]	2.11 [0.12–4.05]	0.07 [0.02–0.11]
2nd quarter (N[d])	53	53	235	235	238
Mean (SE)	5.36 (0.40)	4.85 (0.35)	34.81 (0.52)	50.77 (0.77)	0.70 (0.02)
Difference from baseline [95% CI]	−0.22 [−1.13–0.71]	−0.25 [−1.21–0.60]	−1.08 [−2.43–0.24]	2.32 [0.31–4.25]	0.08 [0.04–0.13]
3th quarter (N[d])	40	40	212	212	221
Mean (SE)	4.95 (0.42)	4.15 (0.39)	35.17 (0.55)	50.89 (0.84)	0.70 (0.02)
Difference from baseline [95% CI]	−0.62 [−1.59–0.31]	−0.95 [−1.80 to −0.10]	−0.72 [−2.17–0.67]	2.44 [0.32–4.48]	0.08 [0.03-0.12]
4th quarter (N[d])	39	39	212	212	208
Mean (SE)	4.37 (0.41)	3.74 (0.38)	34.63 (0.53)	50.91 (84)	0.72 (0.02)
Difference from baseline [95% CI]	−1.21 [−2.17–0.26]	−1.36 [−2.33 to −0.45]	−1.26 [−2.67–0.08]	2.46 [.42–4.54]	0.10 [0.05–0.15]

[a]Higher scores indicate poorer quality of life.
[b]Scores below 50 indicate quality of life is below average of US population.
[c]Lower scores indicate poorer quality of life.
[d]Missing scores were imputed for individuals known to be alive with the posterior predictive distribution of the mean scores at each time point.

participants and those of an average American aged 75 years and over, see Table 7.

Among the 8 health dimensions on the SF-12, bodily pain had the biggest impact on health related quality of life. This was also reflected in the low physical component score observed at baseline, according to which between 12% and 21% of our population were experiencing 'severe' or 'very severe' pain. Interestingly, bodily pain was also the only dimension that showed a negative trend over the whole study period, i.e. reported levels of pain actually increased with respect to those observed at baseline. This trend remained true irrespective of whether individuals were ulcer free and was not bandage related.

The conclusion of an episode of ulceration was associated with an improvement relative to baseline scores on the mental component of the SF-12 and a stabilisation of the physical component, see Table 8. Healed individuals on average experienced an improvement of between two and four points in their MCS with respect to baseline scores. This difference in scores was statistically significant at 6-month and 12-month follow up. The improvement in the MCS suggests that after healing individuals reported to have more frequently had lots of energy, emotional problems caused less disruption in their ability to accomplish and perform daily activities, less frequent episodes of feeling downhearted and low, and an improved ability to carry on their social activities despite physical health and/or emotional problems.

EQ-5D

Utility scores, as measured by the EQ-5D followed a similar a similar pattern to that observed in the SF-12 physical and mental components. After the first 3 months of follow up an improvement in the utility scores was observed; this change in scores of 0.07 with respect to baseline was statistically significant, 95% credibility interval: [0.02–0.11]. Scores became stable around 0.70 thereafter, see Table 6.

Closer examination of the scores indicated that after healing, EQ-5D scores improved on average

Table 7. Comparison with norms-based scores for individuals aged 75 and over

Baseline	Physical component (PSC)	Mental components (MSC)
VenUS study population		
Mean (sd)	35.87 (7.97)	48.50 (12.29)
Range	(4.75–55.27)	(13.11–70.07)
N	339	339
US general population		
Mean (SD)	37.89 (11.16)	50.44 (11.66)
Range	(13–59)	(18–71)
Difference	−2.02	−1.94

21% relative to the scores reported at baseline, see Table 8. Conversely, unhealed individuals over time reported utility scores very close to those reported at baseline. This suggested that the completion of an episode of ulceration was associated with improvements in the degree of mobility, self-care, usual activities, pain/discomfort, and anxiety/depression reported on the EQ-5D questionnaire. Relative to their baseline responses at subsequent follow up contacts, a higher proportion of healed individuals reported no problems with walking about, no problems with self care, no problems with performing usual activities, no pain or discomfort, and no anxiety or depression.

Discriminative properties

Hyland was associated with a small (ES = 0.2) to moderate (ES = 0.6) ability to discriminate individuals according to age, mobility, leg ulcer size, and current ulcer duration, (see Table 7). Unlike the Hyland, the EQ-5D was only associated with a moderate ability to differentiate between individuals in relation to their degree of mobility (ES = −0.7). The SF-12 was associated with small (ES = 0.3) to large (ES = 0.9) ability to differentiate between individuals on the basis of their age, degree of mobility and duration of present ulcer, see Table 9.

Evaluative properties (responsiveness to healing)

The responsiveness of EQ-5D and SF-12 to changes in HRQoL following the ulcer healing at 3 months was confirmed by the ES and SRM estimates. Small to moderate ability to detect changes in individuals' health related quality of life following the closure of their venous leg ulcer, was associated with both SF-12 and EQ-5D at three months, see Table 10. Whilst evidence of an improvement on the scores of healed individuals was observed in seven of the eight individual dimensions described by the SF-12, the bodily pain dimension was associated with a moderate worsening.

Discussion

In the context of the largest randomised controlled trial of compression therapy in the treatment of venous leg ulcers (VenUS I), the health related quality of life of this group of individuals was measured using one disease specific (Hyland) and two generic (SF-12 and EQ-5D) instruments.

The validity and reliability of the Hyland questionnaire was confirmed in a sample of 387 individuals with venous leg ulcers. Unlike the original scale construction, our data suggested a two factor solution for the third section of the Hyland questionnaire. A one-factor solution was associated with poorer discriminatory abilities and responsiveness over time than the two-factor one, detailed results not presented, but available on request from the corresponding author. Factor one appeared to measuring perceptions of more practical limitations associated with the ulcer, whilst factor two was measuring more emotional responses. Descriptive analysis of the practical and emotional factors suggested a worsening in both health related factors associated with venous leg ulceration.

The responsiveness to healing of the Hyland questionnaire is ambiguous. Individuals participating in a pilot study of the VenUS I HRQoL instruments raised concerns regarding the relevance of a number of items in the Hyland questionnaire for individuals without an open venous leg ulcer. The research team considered that unless the Hyland questionnaire was substantially modified the validity of the responses to the Hyland questionnaire after healing could be compromised. Consequently, and to prevent confusion by requesting partial response to a questionnaire after healing and to minimise the burden on the study participants, the team decided to administer the Hyland questionnaire only to individuals with an

Table 8. Mean scores estimates for healed and unhealed individuals[c]

Period	SF-12						EQ-5D		
	Physical component			Mental component			Utility scores		
	Healed[a]	Unhealed[b]	Difference [95% CI*]	Healed[a]	Unhealed[b]	Difference [95% CI]	Healed[a]	Unhealed[b]	Difference [95% CI]
Baseline N		339			339		328		
Mean (SE)		35.88 (0.42)	NA		48.47 (0.64)	NA		0.62 (0.02)	
1st quarter (N)	116	150		116	150		110	151	
Mean (sd)	35.71 (0.72)	34.39 (0.64)	1.41 [−0.46–3.24]	50.49 (1.11)	50.46 (0.98)	0.03 [−2.93–2.94]	0.75 (0.03)	0.64 (0.02)	0.11 [0.04–0.17]
Difference from baseline [95% CI]	−0.17 [−0.79–1.42]	−1.59 [−3.04 to 0.10]		2.02 [−0.50–4.51]	1.97 [−0.33–4.31]		0.13 [0.07–0.20]	0.02 [−0.02–0.08]	
2nd quarter (N)	133	102		133	102		137	101	
Mean (sd)	35.44 (0.68)	33.97 (0.75)	1.47 [−0.58–3.51]	52.65 (1.03)	48.36 (1.20)	4.29 [1.20–7.38]	0.78 (0.03)	0.61 (0.03)	0.16 [0.09–0.23]
Difference from baseline [95% CI]	−0.44 [−2.02–1.14]	−1.91 [−3.62 to 0.21]		4.18 [1.79–6.57]	−0.11 [−2.76–2.55]		0.16 [0.10–0.21]	−0.005 [−0.07–0.06]	
3rd quarter(N)	138	74		138	74		141	80	
Mean (sd)	36.09 (0.67)	33.54 (0.89)	2.55 [0.36–4.81]	51.93 (1.02)	49.22 (1.39)	2.70 [−0.66–6.04]	0.74 (0.02)	0.61 (0.03)	
Difference from baseline [95% CI]	0.22 [−1.35–1.79]	−2.34 [−4.25 to 0.41]		3.45 [1.09–5.82]	0.74 [−2.24–3.75]		0.12 [0.07–0.17]	−0.004 [−0.07–0.06]	0.13 [0.05–0.20]
4th quarter (N)	148	64		148	64		147	61	
Mean (sd)	35.46 (0.64)	32.84 (0.96)	2.62 [0.37–4.92]	52.41 (0.98)	47.66 (1.49)	4.75 [1.21–8.20]	0.75 (0.02)	0.65 (0.04)	0.10 [0.18–0.18]
Difference from baseline [95% CI]	−0.42 [−1.93–1.07]	−3.04[−5.13 to 1.01]		3.94 [1.64–6.22]	−0.81 [−3.99–2.37]		0.13 [0.08–0.19]	0.03 [−0.04–0.11]	

CI = Credibility Interval.

[a]Participants whose ulcer has healed and remains healed at the end of each follow up period.

[b]Participants whose ulcer has not healed or re-occurred within a follow up period.

[c]Missing scores were imputed for individuals to be alive with the posterior predictive distributions of the mean scores at each time point.

Table 9. Effect sizes according to baseline characteristics

	Age[a]		Mobility[b]		Area[c]		Duration[d]	
	ES[e]	*p*	ES	*P*	ES	*p*	ES	*p*
EQ-5D[f]	0.1	0.0	−0.7	0.0	0.0	0.3	0.0	0.2
Hyland[g]								
Practical	0.0	0.2	0.6	0.3	0.2	0.3	0.2	0.2
Emotional	−0.1	0.6	0.3	0.3	0.2	0.3	0.2	0.7
SF−12[f]								
General health	−0.2	0.3	−0.7	0.8	0.0	0.5	−0.1	0.2
Physical functioning	−0.4	0.0	−1.1	0.0	−0.1	0.3	−0.2	0.5
Role physical	−0.2	0.0	−0.9	0.1	−0.2	0.5	−0.1	0.4
Role emotional	−0.1	0.6	−0.4	0.0	−0.1	0.5	0.0	0.7
Body pain	−0.2	0.1	0.5	0.4	0.0	0.5	−0.1	0.5
Mental health	0.0	0.3	−0.3	0.6	0.1	0.9	0.0	0.3
Vitality	−0.2	0.8	−0.5	0.7	−0.1	0.5	−0.2	0.2
Social functioning	−0.2	0.6	−0.8	0.1	−0.2	0.5	−0.1	0.2
Physical score	−0.4	0.0	−0.9	0.2	−0.1	0.9	−0.3	0.5
Mental score	0.0	0.5	−0.4	1.0	−0.1	0.5	0.0	0.4

[a]Cut-off point = sample mean age. Group 1 ⩽71 years; group 2 >71 years.
[b]Mobility: group 1 = fully mobile; group 2 = partially mobile.
[c]Cut-off point = sample median initial ulcer size. Group 1 ⩽4 cm^2 group 2 >4 cm^2.
[d]Cut-off point = median duration of current ulcer. Group 1 ⩽3 months, group 2 > 3 months.
[e]Effect size = (mean in group 2 - mean in group 1)/Overall SD.
[f]Higher scores indicate better quality of life (qol). A positive ES indicate that the qol of individuals in group two is better.
[g]Higher scores indicate poorer quality of life. A positive ES indicates that the qol individuals in group two is poorer.

Table 10. Responsiveness at first 3-month follow up

	Unhealed			Healed		
	N	ES[a]	SMD	N	ES*	SMD
SF-12						
General health	153	−0.1	−0.1	119	0.0	0.1
Physical functioning	155	0.0	0.0	119	0.0	0.1
Role physical	151	0.2	0.2	118	0.3	0.3
Role emotional	158	−0.1	−0.1	124	0.2	0.2
Body pain	153	−0.3	−0.3	125	−0.5	−0.5
Mental health	154	0.1	0.1	126	0.1	0.1
Vitality	155	0.0	0.0	124	0.0	0.0
Social functioning	154	0.2	0.2	125	0.1	0.2
Physical score	140	−0.2	−0.2	113	−0.2	−0.2
Mental score	140	0.1	0.1	113	0.2	0.3
EQ-5D	134	0.1	0.1	108	0.5	0.4

*Negative values indicate a decline in HRQoL with respect to baseline values.

open ulcer. Further research work is required to warrant the responsiveness to venous leg ulcer healing of the Hyland questionnaire.

Our results coincide with those of Walters et al. [8] and Franks et al. [9] i.e. in that generic measures (SF-12 and EQ-5D) have poor abilities to discriminate individuals according to baseline ulcer characteristics; no unexpected given the generic nature of these instruments. In contrast to the findings of Walters et al. [8], and similarly to those from Franks et al. [9] our data indicated that the two generic instruments *were* responsive to changes in HRQoL in relation to ulcer status (healed/unhealed), irrespective of the follow up

period. On average individuals whose venous leg ulcer healed and remained so within the first year of follow up reported higher SF-12 and EQ-5D scores than those whose ulcer remained unhealed. This suggests that the resolution of an episode of open ulceration results in an improvement on the HRQoL of venous leg ulcer patients. Differences in the time points for follow up could account for the inconsistencies in the results from these three studies. The VenUS I is the only study in which quarterly measurements were made.

A number of studies have highlighted the impact of pain in people with venous leg ulcers [5, 6, 9]. In this analysis high levels of bodily pain were detected by the SF-12 irrespective of whether ulcers were healed or unhealed. Conversely, levels of pain specifically associated with the presence of an open venous leg ulcer reported in the Hyland questionnaire were moderate; in the first year of follow up 50% of individuals indicated that their ulcer only 'hurt a little'. This suggests that the high levels of bodily pain reported by this population are likely to be primarily due to something other than their open venous leg ulcer. The exploration of the contribution of alternative co-morbidities to the high levels of bodily pain was prevented by the unavailability of such data from the VenUS I study population. Future venous leg ulcer studies should collect data which differentiates between ulcer related and unrelated pain, and attempt to determine the sources of pain unrelated to ulceration.

The Hyland questionnaire was associated with the lowest partial and complete response rates 49% at 1 year of follow up. Response rates to the EQ-5D and SF-12 also reduced over time to 59% and 61% at 1 year, respectively. A number of measures were implemented to improve response rates during the study's follow up period. For example, non-respondent participants were sent monthly reminders accompanied by a new copy of the quality of life questionnaire. Only individuals who indicated their unwillingness to provide any further information were taken off the mailing list. Participants were not contacted by phone since pooling health related quality of life data reported at two different points in time may result in inconsistencies with the data.

The relevance of disease specific instruments to the acquisition of a deeper understanding of the impact of different health care conditions in HRQoL is widely recognised. In the area of venous leg ulceration such an instrument would need to consider the chronic nature of venous leg ulcers. In other words, it will have to describe the main aspects/dimensions of HRQoL that are altered as consequence of the open ulcer, as well as from its underlying causes, and consider both the impact of fear of recurrence and measures taken to prevent recurrence. Bearing that in mind, it is important to recognise that venous leg ulcers are mainly present among elderly people. This population group is likely to be affected by co-morbidities. The ability of a venous leg ulcer disease specific instrument to disentangle the changes in HRQoL solely attributable to the open ulcer, and/or its underlying causes is unclear. For example; CIVIQ [11] and Frieburg [12] questionnaires are both health measures for venous insufficiency, however this condition covers a spectrum of disorders ranging from varicose veins through to extensive ulceration of the leg, thus it can not be argued that the changes in health related quality of life captured by these instruments are principally related to venous ulceration and/or its underlying causes.

While disease specific instruments can provide a deeper insight into the effect of different health conditions on health related quality of life, this is usually done at the cost of increasing questionnaire's length. The three existing disease specific instruments available in the area of venous leg ulcers – Frieburg, CIVIQ and Hyland – are considerably longer than the generic instruments used in this study (SF-12 and EQ-5D). In the context of randomised controlled studies where one of the main objectives is to minimise potential sources of bias, there is always a trade off between the quantity and the quality of information to be collected. Previously published studies indicate that by decreasing the overall length of quality of life questionnaire booklets, response rates and quality of responses may be improved [32].

In conclusion, this study suggests that HRQoL in individuals with venous leg ulcers can be effectively measured with generic instruments. Evidence of the impact of venous leg ulceration on health related quality of life was provided by both generic and disease specific instruments. The disease specific Hyland questionnaire had good

discriminative abilities, however it was associated with considerably poorer response rates than the SF-12 and the EQ-5D. Moreover, the responsiveness to healing of the Hyland questionnaire is unclear. The two generic instruments, SF-12 and EQ-5D, performed reasonably well, both had response rates between 93% and 54% and were responsive to changes in health related quality of life occurring after the closure of a venous leg ulcer. Consequently, in the context of randomised controlled trials in the area of chronic venous leg ulcers, we recommend the use of generic measurements of HRQoL.

Acknowledgement

We thank the participants for taking part in the trial; district nurses and hospital out-patients staff for recruiting patients into the trial and completing trial documentation; research nurses at the initial four centres (Una Adderley, Sharon Hailes, Nikki Stubbs, and Angela Williams), and local investigators at satellite sites (Jane Best, Sian Coupe, Caroline Dowsett, Val Duncan, June Jones, Robert Smith and Gerry Young) for coordinating local patient recruitment and helping ensure follow-up data was returned.

The VenUS I collaborators (current and *past*) are; *Peter Anderson*, David Berridge, Nicky Cullum, Alun Davies, Amanda Farrin, Andrew Garratt, Lesley Godfrey, Cynthia Iglesias, Andrea Nelson, Diane Reddington, Julian Scott, Mike Walker. Una Adderley, David Berridge, Alun Davies, Andrew Garratt, June Jones, Julian Scott, Nikki Stubbs, Elizabeth Scanlon.

Contributors: YB performed and interpreted the factor analysis. NAC was the principal investigator, contributed to the study design, coordinated the project, and contributed to writing the manuscript. CI was the data manager, health economist, conducted the statistical analysis of the HRQoL data analysis and was the main responsible for writing up the manuscript. EAN was the clinical trial coordinator, contributed to the study design and commented on previous versions of the manuscript. Elizabeth Scanlon contributed to the study design and commented on previous versions of the manuscript. NAC will act as guarantor for the paper.

We would also like to thank three anonymous referees and Elaine McColl for their insightful comments on previous versions of this manuscript and Andrea Manca for his advise on multilevel modelling. Sources of support: This trial was funded by the UK National Health Service Research and Development HTA Programme but the views and opinions expressed do not necessarily reflect those of the NHS Executive.

Competing interests: none.

References

1. Chase SK, Melloni M, Savage A. A forever healing: The lived experience of venous ulcer disease. J Vasc Nurs 1997; 15: 73–78.
2. Walshe C. Living with a venous leg ulcer: A descriptive study of patients' experiences. J Adv Nurs 1995; 22: 1092–1100.
3. Ebbeskog B, Ekman SL. Elderly persons' experiences of living with venous leg ulcer: Living in a dialectal relationship between freedom and imprisonment. Scand J Caring Sci 2001; 15: 235–2343.
4. Hyde C, Ward B, Horsfall J, Winder G. Older women's experience of living with chronic leg ulceration. Int J Nurs Pract 1999; 5: 189–198.
5. Flett R, Harcourt B, Alpass F. Psychosocial aspects of chronic lower leg ulceration in the elderly. West J Nurs Res 1994; 16: 183–192.
6. Roe B, Cullum N, Hamer C. Patients' perceptions of chronic leg ulceration. In: Leg Ulcers; Nursing Management. A Research Based Guide. London: Balliere Tindall, 1998: 125–134.
7. Wissing U, Ek AC, Unosson M. A follow-up study of ulcer healing, nutrition, and life-situation in elderly patients with leg ulcers. J Nutr Health Aging 2001; 5: 37–42.
8. Walters SJ, Morrell CJ, Dixon S. Measuring health-related quality of life in patients with venous leg ulcers. Qual Life Res 1999; 8: 327–336.
9. Franks PJ, Moffatt CJ. Health related quality of life in patients with venous ulceration: Use of the Nottingham health profile. Qual Life Res 2001; 10: 693–700.
10. Hyland M. Quality of life of leg ulcer patients: questionnaire and preliminary findings. J Wound Care 1994; 3: 294–298.
11. Launois R, Reboul-Marty J, Henry B. Construction and validation of a quality of life questionnaire in chronic lower limb venous insufficiency (CIVIQ). Qual Life Res 1996; 5: 539–554.
12. Augustin M, Dieterle W, Zschocke I, et al. Development and validation of a disease-specific questionnaire on the quality of life of patients with chronic venous insufficiency. Vasa 1997; 26: 291–301.
13. Ware J, Jr, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Med Care 1996; 34: 220–33.
14. Kind P. Quality of Life and Pharmacoeconomics in Clinical Trials. Philadelphia: Lippincott–Raven, 1996.

15. Hyland M. A brief gruide to the selection of quality of life instrument. Health Qual Life Outcomes 2003; 1: 24–28.
16. Norusis M. SPSS for windows (Version 10.0), 1986.
17. Cattell R. The scree test for the number of factors. Multivar Behavi Res 1966; 1: 245–276.
18. Kline P. Psychology Exposed: Or, The Emperor's New Clothes. London: Routledge, 1988.
19. Friendly M. Planning a factor analytic study, 2003. http://www.psych.yorku.ca/lab/psy6140/fa/facplan.htm.
20. Cronbach LJ. Coefficient alpha and the internal structure of tests. Psycometrika 1951; 16: 297–334.
21. Kind P, Hardman G, Macran S. UK Population Norms for EQ-5D. 172. Centre for Health Economics Discussion Paper, University of York, 1999.
22. Iglesias CP, Birks YF, Torgerson DJ. Improving the measurement of quality of life in older people: The York SF-12. QJM 2001; 94: 695–698.
23. Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. J Clin Epidemiol 2003; 56: 395–407.
24. Fairclough DL. Design and Analysis of Quality of Life Studies in Clinical Trials. Chapman & Hall/CRC, 2002.
25. Goldstein H, Browne W, Rasbash J. Multilevel modelling of medical data. Stat Med 2002; 21: 3291–315.
26. Goldstein H, Healy NJR, Rasbash J. Multilevel time series models with applications to repeated measures data. Stat Med 1994; 13: 1643–1655.
27. Spiegelhalter DJ, Thomas A, Best NG. WinBUGS Version 14. MRC Biostatistics Research Unit Cambridge, 2003.
28. Gelman A. Prior distributions for variance parameters in hierachical models. In: Econometrics from Economics Working Paper Archive at WUSTL, 2004.
29. Abrams KR, Brazier J, O'Hagan T, Kharoubi S, Tsuchiya A. Modelling partially & completely missing preference based outcome measures (PBOMs). http://www shef ac uk/chebs/FFortnights/FF6/Abrams ppt 2003.
30. Iglesias CP, Nelson EA, Cullum N, Torgerson DJ. VenUS I: A randomized controlled trial of two types of bandage for treating venous leg ulcers. Health Technol Assess NHS R&D HTA Program 2004; 8.
31. Ware JE, Kosinski M. SF-36 Physical and mental health summary scales: A manual for users of Version 1.2., 2001.
32. Iglesias C, Torgerson DJ. Does length of questionnaire matter? A randomised controlled trial of response rates to a mailed questionnaire. J Health Serv Res Policy 2000; 5: 219–221.

Address for correspondence: C.P. Iglesias, Department of Health Sciences, Seebohn Rowntree Building, Area 2, Room 148, York, North Yorkshire, Y010 5DD, UK
E-mail: cpiu1@york.ac.uk

Randomized trial

Randomized clinical trial of four-layer and short-stretch compression bandages for venous leg ulcers (VenUS I)

E. A. Nelson, C. P. Iglesias, N. Cullum and D. J. Torgerson on behalf of the VenUS I collaborators

Correspondence to: Professor N. Cullum, Department of Health Sciences, Seebohm Rowntree Building (Area 4), University of York, Heslington, York YO10 5DD, UK (e-mail: nac2@york.ac.uk)

Background: A randomized clinical trial was undertaken to determine the relative effectiveness of four-layer and short-stretch bandaging for venous ulceration.

Methods: A total of 387 adults with a venous ulcer, who were receiving leg ulcer treatment either in primary care or as a hospital outpatient, were recruited to this parallel-group open study and randomized to either four-layer or short-stretch bandages. Follow-up continued until the patient's reference leg was ulcer free or for a minimum of 12 months. The primary endpoint was time to complete healing of all ulcers on the reference leg. Secondary outcomes included proportion of ulcers healed, health-related quality of life, withdrawals and adverse events. Analysis was by intention to treat.

Results: Unadjusted analysis identified no statistically significant difference in median time to healing: 92 days for four-layer and 126 days for short-stretch bandages. However, when prognostic factors were included in a Cox proportional hazards regression model, ulcers treated with the short-stretch bandage had a lower probability of healing than those treated with the four-layer bandage: hazard ratio 0·72 (95 per cent confidence interval 0·57 to 0·91). More adverse events and withdrawals were reported with the short-stretch bandage.

Conclusion: Venous leg ulcers treated using a four-layer bandage healed more quickly than those treated with a short-stretch bandage.

Presented to the Annual Meeting of the Vascular Surgery Society of Great Britain and Ireland, Glasgow, UK, November 2003. Winner of the Richard Wood prize and joint winner of the Venous Forum prize

Paper accepted 18 May 2004

Published online in Wiley InterScience (www.bjs.co.uk). **DOI:** 10.1002/bjs.4754

Introduction

Leg ulceration affects 15–18 per 1000 adults in developed countries, and is associated with pain and lower quality of life[1,2]. The majority of leg ulcers are secondary to venous insufficiency (venous ulcers). It has been estimated that the management of venous ulcers in the UK costs £100–300 million every year, nursing time accounting for most of this cost[3,4].

Systematic reviews have found that compression therapy heals more ulcers than dressings alone, and that high-level compression systems are more effective than low- or moderate- compression systems[5]. However, it has not been possible to determine from previous research which high-compression regimen – four-layer bandaging, compression hosiery, Unna's boot or short-stretch bandaging – is the most clinically or cost effective[5]. Four-layer bandaging is probably the most widely used method in the UK, whereas short-stretch bandaging is the system of choice in much of continental Europe. Thus four-layer and short-stretch bandages were chosen for comparison. Short-stretch bandage is washable and reusable (unlike four-layer bandages), comparatively simple to apply, and its effects are thought to be less reliant on the skill of the bandager[6]. Furthermore, it is inelastic, a feature that may reduce the likelihood of damage to the skin and underlying tissue occasionally induced by inappropriately applied elastic high compression. To determine the relative effectiveness and cost-effectiveness of four-layer and short-stretch bandaging systems, a pragmatic randomized clinical trial was undertaken.

Patients and methods

Patients were eligible for the trial if they had been diagnosed clinically with a venous leg ulcer at least

1 cm in diameter. Exclusion criteria were: age less than 18 years, significant arterial disease (ankle : brachial pressure index (ABPI) less than 0·8), diabetes mellitus, previous unsuccessful use of a trial bandage, patient unable or unwilling to have high compression, and patient unable or unwilling to provide written, informed consent.

Between April 1999 and December 2000, patients with venous leg ulcers treated in the community (leg ulcer services, district nursing or general practice) or as an outpatient (vascular surgery) were recruited from nine UK centres. After obtaining written informed consent, participants were allocated to either four-layer bandages (4LB group) or short-stretch bandages (SSB group) using a concealed, remote telephone randomization service. Randomization was stratified by clinical centre, previous ulceration (yes or no), ulcer area (less than or greater than 10 cm^2), and ulcer duration (more or less than 6 months)[7]. The randomization code was developed using computer-generated permuted blocks (randomly sized 4 or 6). Both patients and nurses were aware of the allocated treatment after assignment.

Bandaging

Four-layer bandages consisted of one layer of orthopaedic wool padding, covered by three retention/compression bandages, all applied with 50 per cent overlap (*Table 1*). The original four-layer bandage system and two proprietary kits, Profore® (Smith and Nephew Healthcare, Hull, UK) and System 4® (SSL International, Knutsford, UK), both 10 cm wide, were randomly allocated in this pragmatic trial as a previous study that compared these systems found no significant difference in effectiveness[8]. Standard modifications were used for participants with large (circumference greater than 25 cm) or small (circumference less than 18 cm) ankles (*Table 1*)[9]. Bandages were used once and discarded.

Table 1 Four-layer bandage system and method of application

Layer	Type of bandage (ankle circumference 18–25 cm)	Method of application
1	Padding (e.g. Velband®, Johnson & Johnson)	Spiral
2	Retention (e.g. Crepe® BPI)	Spiral
3	Class 3A compression (e.g. Elset®, SSL International)	Figure of eight
4	Cohesive compression (e.g. Coban®, 3M)	Spiral

Two layers of padding (layer 1) were used for participants with an ankle circumference of less than 18 cm. The third bandage layer was replaced with a high-compression (class 3C) bandage for participants with an ankle circumference greater than 25 cm.

Short-stretch bandaging comprised one layer of orthopaedic wool padding covered with one or two compression bandages, either Comprilan® (Beiersdorf UK, Milton Keynes, UK) or Rosidal K® (Vernon-Carus, Preston, UK) 100 per cent cotton short-stretch bandages. The standard techniques of spiral, figure of eight or modified Putter were used[10]. The different techniques are thought to deliver similar sub-bandage pressures. The bandages were, whenever possible, washed by the patient and reused.

In both groups the ulcers were cleansed using tap water or saline, and covered with a simple, low-adherent dressing, such as N-A Dressing® (Johnson & Johnson Medical, Ascot, UK). Dressings and bandages were renewed by the regular nursing staff; any changes to dressings and the frequency of renewal were decided by the patient's usual nurse. As this was a pragmatic trial, the schedule of patient visits was not dictated by trial protocol, except to say that patients should be seen at least once a week.

Data collection

Nurses completed a dressing log at each leg ulcer dressing visit, which recorded whether or not an ulcer was healed, the date of the visit, all materials used, and reasons for any changes in treatment. Health-related quality of life was assessed by means of Short Form SF-12[11], EuroQol-5D[12], and Hyland leg and foot ulcer questionnaire[13]. Economic data were also recorded and are reported elsewhere[14]. Follow-up of all participants continued from randomization until December 2001 (12–21 months).

The primary endpoint was time to healing of all ulcers on the reference leg, defined as the leg with the largest eligible ulcer. Secondary outcomes included the proportion of ulcers healed at 12 and 24 weeks, withdrawals and adverse events. A healed ulcer was defined as complete epithelial cover in the absence of a scab. At healing, a nurse took a Polaroid™ photograph of the ulcer and sent this to the trial office where an investigator blinded to bandage allocation confirmed ulcer healing. Standard methods of photography were used throughout the study centres to reduce error[15].

Sample size calculation and statistical analysis

Because previous studies had not reported time to event data, the required sample size was estimated using data for the proportion of patients healed at 12 weeks. This was viewed as a conservative approach in the light of a planned survival analysis. The sample size calculation was based on an estimate from previous trials of 50 per cent of ulcers healing by 12 weeks with the four-layer bandage, and an

absolute increase in healing of 15 per cent at 12 weeks using short-stretch bandage. The required sample size (80 per cent power, $\alpha = 5$ per cent) was therefore 400 (200 patients in each arm)[5].

The primary analysis was by intention to treat and the time to complete healing of all ulcers on the reference leg was compared. Time to censoring was defined as the time point at which the follow-up of a patient ceased without the ulcer healing. The primary analysis was a simple comparison of the healing rates using Kaplan–Meier survival curves for the two bandage groups; differences between the two curves were tested by means of the log rank test[16]. To investigate the effects of previously identified prognostic factors for ulcer healing, a Cox regression model was fitted to the data on time to complete healing. Potential prognostic factors considered were: treatment centre, ulcer area, ulcer duration, ulcer episode, age, weight, mobility, ankle mobility and ABPI. Identification of the variables

Table 2 Baseline characteristics of patients and ulcers

	SSB	4LB
Patient characteristics		
Age (years)*	71·3(14·1) (23–96)	71·9(12·3) (25–97)
Height (m)*	1·68(0·11) (1·42–1·96)	1·69(0·11) (1·45–2·03)
Weight (kg)*	79·0(20·3) (38·1–142·4)	80·6(19·3) (33·1–139·7)
Sex ratio (M : F)	80 : 112	79 : 116
Ulcer location		
Left leg	82 (42·7)	88 (45·1)
Right leg	108 (56·3)	107 (54·9)
Patient mobility		
Fully mobile	115 (59·9)	123 (63·1)
Walks with assistance	70 (36·5)	72 (36·9)
Immobile	3 (1·6)	0 (0)
Ankle mobility		
Full	128 (66·7)	131 (67·2)
Reduced	58 (30·2)	59 (30·3)
None (fixed)	2 (1·0)	3 (1·5)
Short form 12 score*		
Physical component	35·6(7·6) (18·3–52·4)	36·1(8·3) (4·7–55·3)
Mental component	48·4(12·3) (16·8–68·4)	48·6(12·3) (13·1–70)
Leg and ulcer characteristics		
Time since onset of first ulcer (years)†	4 (0–75)	3 (0–60)
No. of episodes (since onset)†	2 (0–64)	2 (0–50)
Ankle circumference (cm)	23·9(2·9) (16·0–32·3)	23·9(2·9) (16·2–34·0)
Duration (months)†	3 (0–768)	3 (0–456)
Area (cm^2)†	3·82 (0·35–143·93)	3·81 (0·19–254·58)
Ankle : brachial pressure index	1·04(0·14) (0·80–1·62)	1·05(0·15) (0·80–1·90)
Ulcer type		
Sloughy	108 (56·3)	130 (66·7)
Granulating	115 (59·9)	127 (65·1)
Epithelializing	25 (13·0)	27 (13·8)
Skin features		
Eczematous	49 (25·5)	59 (30·3)
Macerated	26 (13·5)	31 (15·9)
Cellulitis	15 (7·8)	15 (7·7)
Lipodermatosclerosis	86 (44·8)	88 (45·1)
Stratifying variables		
Ulcer area (cm^2)		
≤ 10	158 (82·3)	158 (81·0)
> 10	34 (17·7)	37 (19·0)
Previous ulcer on the trial leg		
Yes	114 (59·4)	115 (59·0)
No	78 (40·6)	80 (41·0)
Ulcer duration (months)		
≤ 6	143 (74·5)	138 (70·8)
> 6	49 (25·5)	57 (29·2)

Values in parentheses are percentages unless indicated otherwise; *values are mean(s.d.) (range); †values are median (range). 4LB, four-layer bandages; SSB, short-stretch bandages.

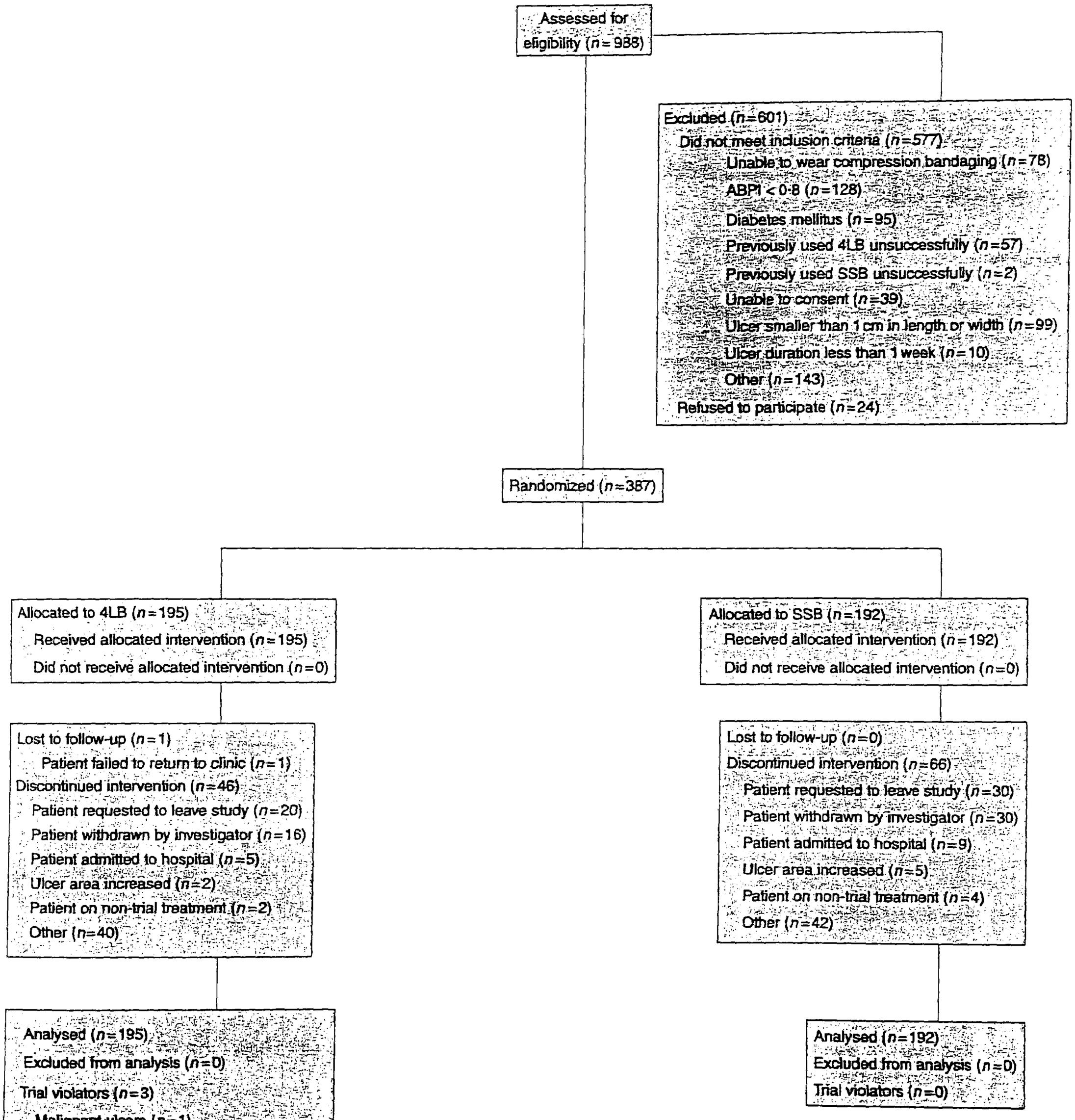


Fig. 1 Flow chart for VenUS I trial. 4LB, four-layer bandages; SSB, short-stretch bandages; ABPI, ankle : brachial pressure index

that best described the hazard of healing the leg ulcers was performed using a procedure described by Collett[17], in which the difference between the −2log likelihood statistic associated with the models being compared was used as the criterion of best model fitting. For completeness of the analysis interactions were investigated, but the trial was not designed to test for these. The assumption of proportionality of hazards between treatment groups

was checked using log cumulative hazard plots, that is $\log(-\log(S(t))$ plotted against $\log t$, where $S(t)$ values were the Kaplan–Meier estimates of the cumulative distribution of healing times. Missing data on demographic and clinical variables considered in the Cox model were inputed using the mean/mode. The largest number of inputed values for any single variable was 12 (3·0 per cent) and the minimum was one (0·3 per cent).

Results

Of 988 patients assessed for trial eligibility, 387 were recruited (*Fig. 1*). Baseline descriptive data were similar for both groups (*Table 2*). *Table 3* shows recruitment by centre.

Time to healing

The median time to ulcer healing in the 4LB group was 92 (95 per cent confidence interval (c.i.) 71 to 113) days, compared with 126 (95 per cent c.i. 95 to 157) days in the SSB group. *Fig. 2* shows the unadjusted Kaplan–Meier curves. The difference in the distribution of cumulative healing times between the two groups was not statistically significant (log rank $\chi^2 = 2{\cdot}46$, $P = 0{\cdot}117$).

Hazard function

As well as including type of treatment, the final best model for the hazard of healing also included treatment centre, number of previous episodes, weight, natural logarithm of baseline ulcer area, ulcer duration and ankle mobility. Estimated coefficients for each of the explanatory co-variates and for the overall adjusted treatment effect are shown in *Table 4*. Having adjusted for these explanatory variables there was a statistically significant increase in the probability of healing in the 4LB group: hazard ratio 0·72 (95 per cent c.i. 0·57 to 0·91).

For completeness of the analysis, interactions between treatment and the six relevant co-variates in the model were investigated. The only interaction that statistically significantly reduced the −2log likelihood statistic (by 18·99) was that between treatment and centre ($P < 0{\cdot}015$). However, given the small number of patients in some of the centres, a further analysis was conducted on only the four centres with the largest number of patients, namely Cumbria, Leeds, West London and North Yorkshire. The centre effect was no longer observed in this latter analysis.

Proportion of ulcers healed

Most previous trials of compression bandaging have not described a survival analysis but reported merely the proportion of ulcers healed at 12 or 24 weeks. Corresponding data for the present study are shown in *Table 5*.

Table 3 Recruitment at each of the trial sites by treatment group

Centre	4LB	SSB
Cumbria	35 (17·9)	35 (18·2)
Leeds	54 (27·7)	52 (27·1)
West London	18 (9·2)	20 (10·4)
North Yorkshire	53 (27·2)	55 (28·6)
Falkirk	9 (4·6)	6 (3·1)
Calderdale	4 (2·1)	5 (2·6)
East London	7 (3·6)	2 (1·0)
Newmarket	4 (2·1)	4 (2·1)
Southport	11 (5·6)	13 (6·8)
Total	195	192

Values in parentheses are percentages. 4LB, four-layer bandages; SSB, short-stretch bandages.

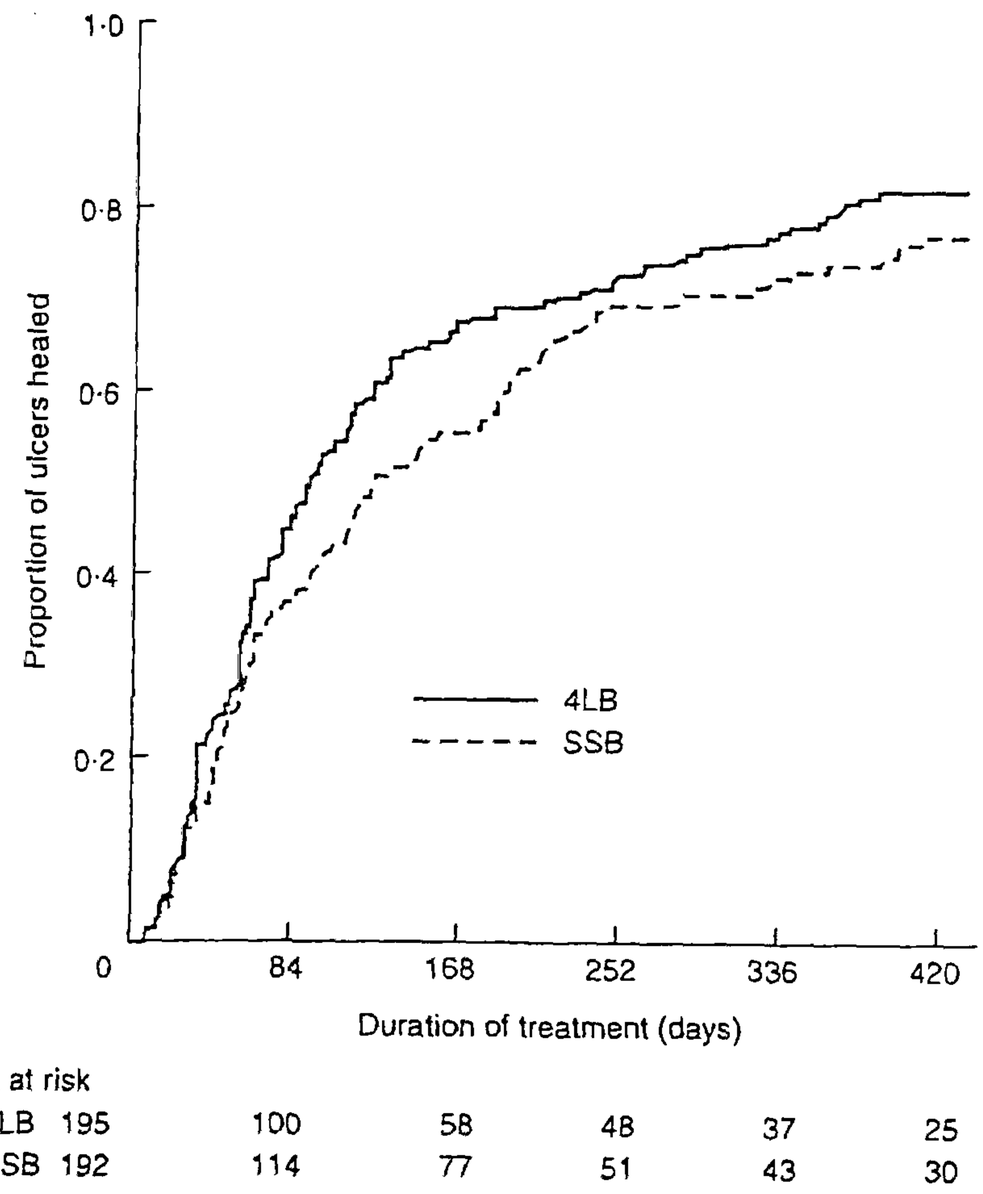


Fig. 2 Kaplan–Meier analysis showing proportion of ulcers healed by treatment with four-layer (4LB) and short-stretch (SSB) bandages

Withdrawals and adverse events

A total of 112 patients withdrew from their original bandage allocation (*Fig. 1*), 46 in the 4LB group and 66 in the SSB group. Of these 20 (4LB) and 30 (SSB) were patient-initiated withdrawals (*Fig. 1*). There were also

Table 4 Treatment effect adjusted for explanatory co-variates

	β_i	Standard error	Exp (β_i)
Bandage (SSB *versus* 4LB)	−0·33	0·12	0·72 (0·57, 0·91)
Previous episodes	−0·04	0·02	0·97 (0·94, 1·00)
Weight	−0·01	0·00	0·99 (0·99, 1·00)
Ulcer area	−0·30	0·06	0·74 (0·66, 0·83)
Duration	−0·02	0·01	0·98 (0·97, 0·99)
Ankle mobility	0·43	0·14	1·53 (1·17, 2·00)
North Yorkshire			1
Leeds	−0·00	0·16	0·10 (0·73, 1·37)
Cumbria	0·70	0·17	2·00 (1·43, 2·82)
West London	0·38	0·23	1·46 (0·93, 2·29)
Southport	−0·10	0·25	0·90 (0·55, 1·47)
Falkirk	−0·17	0·33	0·85 (0·45, 1·60)
Calderdale	0·41	0·38	1·50 (0·72, 3·16)
East London	−0·63	0·43	0·53 (0·23, 1·25)
Newmarket	−0·13	0·52	0·88 (0·32, 2·43)

Values in parentheses are 95 per cent confidence intervals. The sign of each regression coefficient (β_i) gives an indication of whether the hazard increases (positive sign) or decreases (negative sign) for subjects with higher values of that variable. For continuous co-variates the estimated regression coefficient refers to the increase in the log hazard for an increase of 1 in the value of the co-variate. SSB, short-stretch bandages; 4LB, four-layer bandages.

Table 5 Proportion of legs healed during trial

	Proportion of legs ulcer free (%)		
	4LB	SSB	Difference (%)
12 weeks	46·3	36·7	9·6 (0, 20)*
24 weeks	67·5	55·4	12·1 (2, 22)†

Values in parentheses are 95 per cent confidence intervals. 4LB, four-layer bandages; SSB, short-stretch bandages. *$P = 0·1$. †$P = 0·02$. *T*-test.

Table 6 Adverse events possibly related to compression treatment

	4LB	SSB	Total
Maceration	18	17	35
Excoriation	10	14	24
Skin damage	49	55	104
Bandage failure	37	36	73
Ulcer deterioration (including infection)	91	166	257
Skin deterioration	30	27	57
Surgical intervention to leg	4	5	9
Dryness	2	1	3
Non-surgical hospitalization related to leg ulceration	2	0	2
Occurrence of new ulcer	11	13	24
Medical event relating to leg	1	3	4
Total no. of adverse events	255	337	592
Total no. of patients with adverse events	76	91	167

4LB, four-layer bandages; SSB, short-stretch bandages.

Table 7 Healing rates at various time points in trials comparing four-layer and short-stretch bandages

		Proportion of ulcers healed (%)					
		12 weeks		16 weeks		52 weeks	
Reference	Year	4LB	SSB	4LB	SSB	4LB	SSB
Duby *et al.*[19]*	1993	44	40	–	–	–	–
Scriven *et al.*[20]*	1998	34	41	–	–	53	56
Partsch *et al.*[21]†	2001	–	–	62	73	–	–
Ukat *et al.*[22]	2003	30	22	–	–	–	–
VenUS I†	2004	46	37	55	45	78	72

*In these trials the legs of several patients with bilateral ulceration were randomized independently. †Primary outcome in these trials was complete healing of all ulcers on trial leg (in the presence of multiple ulcers this would tend to underestimate the outcome of ulcer healing). 4LB, four-layer bandages; SSB, short-stretch bandages.

significantly more nurse-initiated withdrawals in the SSB group (30 *versus* 16; $P = 0·035$). A higher proportion of patients experienced non-bandage-related adverse events, such as health problems, change in ulcer diagnosis or death, in the SSB group (39 *versus* 33), but this was not statistically significant ($P = 0·468$).

Overall, 592 adverse events in 167 patients were classified as possibly bandage related (*Table 6*). More patients had clinically diagnosed infected ulcers in the SSB group than in the 4LB group (46 and 32 respectively) but this was not statistically significant ($\chi^2 = 3$, 1 d.f., $P = 0·084$). In the majority of these patients it was the clinical impression of the nurse that the ulcer was infected; ulcers were not routinely swabbed for infection as there is no consensus on how to define or detect infection in chronic wounds.

Discussion

This trial demonstrated that venous ulcers treated with four-layer bandages were significantly more likely to heal than those treated with short-stretch bandages; the median time to healing was 34 days less with four-layer bandages. There was a statistically significant interaction between bandage and centre, but there was no reason to suspect a centre effect beforehand and so this had not been planned as part of the primary analysis[18]. The reasons for this interaction are unclear. However, heterogeneity in the results was no longer observed when the centres that recruited the smallest number of patients were excluded from the analysis, suggesting that the interaction effect was more likely to be a chance finding.

Real centre differences relating to nursing factors such as previous bandage experience and bandage skill cannot be discounted. During the trial set-up phase, training in the application of both bandages was provided as required. Nurses were more likely to have previous training and experience of four-layer bandaging, although it only became prescribable in primary care shortly before the trial commenced. Future trials should assess the bandage application skills of participating nurses at baseline, and either reduce the heterogeneity in skill through training or analyse the effects of these differences.

Although fewer ulcers deteriorated during four-layer bandaging and there were more nurse-defined ulcer infections with short-stretch bandaging, the differences between groups were not statistically significant. However, the possibility of a real difference in infection rate cannot be excluded as the trial was underpowered for this outcome. The lack of agreed and objective criteria for defining ulcer infection in chronic wounds (which are usually colonized by bacteria but not infected) means that nurses relied on their clinical impression. Patients in the SSB group were at greater risk of infection as their ulcers were open for longer, and the (unblinded) nurses applying short-stretch bandaging may have been more vigilant in the detection and reporting of infection. The significantly higher number of nurse-initiated withdrawals in the SSB group may also have been related to the lack of blinding of nurses and patients.

Unlike most other wound care trials, time to healing was the main outcome measure, rather than proportion healed at a specific time point. Any choice of a discrete time point is largely arbitrary, and the outcome may depend on the time point chosen. Furthermore, patients are more likely to want to know how long their ulcer will take to heal than whether it is likely to be healed in a certain number of weeks. In this trial there was no significant difference in the proportion of ulcers healed at 12 weeks, but a statistically significant difference at 24 weeks. The fact that the primary, unadjusted time to healing analysis did not identify a significant difference between the bandages, whereas the Cox regression analysis determined the difference to be statistically significant, is worthy of comment. It is standard practice to adjust the analysis for factors deemed likely to influence healing (they are also stratifying variables at randomization). Minor imbalances between treatment groups at baseline for factors such as ulcer duration can affect the precision of the estimate of the difference in healing times, and may lead to the difference not being statistically significant. In this study, adjustment of the analysis for prognostic factors did not reverse the main result, but merely increased the precision of the estimates, with the difference becoming statistically significant[18].

Four other randomized clinical trials have compared short-stretch and four-layer bandages, two of which were published after this trial commenced recruitment (*Table 7*)[19–22]. All of these trials were small, involving 317 participants in total (compared with 387 in this trial). Only the largest trial by Partsch *et al.*[21] conducted a survival analysis, but no statistically significant difference in ulcer healing between the bandages was shown. The other studies reported crude proportions healed at 3 or 4 months. The present study followed patients for a minimum of 12 months and captured the healing of most ulcers.

The healing rates achieved in this trial were similar to those reported in the four previous studies (*Table 7*). Patient eligibility criteria for the five trials were broadly similar. The trial by Ukat *et al.*[22] included only ulcers with a baseline area greater than 4 cm^2, which may in part account for the lower rate of ulcer healing at 12 weeks than in the other trials. Of studies that reported proportions of ulcers healed, with no adjustment for baseline differences, those by Duby *et al.*[19] and Scriven *et al.*[20] both had an imbalance in baseline ulcer area that may have contributed to the final differences in healing rates between the bandages. The average ulcer size in the study of Partsch *et al.*[21] was lower than that in the other trials, which partly explained their superior healing rates at 16 weeks.

Combining data from the VenUS I and Partsch[21] studies resulted in a pooled hazard ratio that favoured four-layer bandaging, although this was not significant: hazard ratio 0·85 (95 per cent confidence interval 0·49 to 1·48). The two trials may in fact have investigated the impact of the *introduction* of a high-level compression bandage system into a setting where an acceptable mode of compression was already in use. A balanced conclusion could be that, where four-layer bandages are currently in use, they should remain the system of choice. However, short-stretch bandaging may be a useful alternative for patients who do not tolerate the four-layer system. Where short-stretch bandages are already use, an audit of bandage renewal frequency and healing rates should be conducted to ascertain whether such bandaging is cost effective[14].

Acknowledgements

The VenUS I collaborators (past and present) are: Una Adderley, Peter Anderson, David Berridge, Nicky Cullum, Alun Davies, Amanda Farrin, Andrew Garratt, Lesley Godfrey, Cynthia Iglesias, June Jones, Andrea Nelson, Diane Reddington, Julian Scott, Nikki Stubbs, David Torgeson and Mike Walker.

This trial was funded by the UK National Health Service (NHS) Research and Development HTA Programme but the views and opinions expressed do not necessarily reflect those of the NHS Executive. Trial-related education was provided by Beiersdorf, a supplier of short-stretch bandages.

The authors thank the participants for taking part in the trial; district nurses and hospital outpatient staff for recruiting patients into the trial and completing trial documentation; research nurses at the initial four centres (Una Adderley, Sharon Hailes, Nikki Stubbs and Angela Williams) and local investigators at satellite sites (Jane Best, Sian Coupe, Caroline Dowsett, Val Duncan, June Jones, Robert Smith and Gerry Young) for coordinating local patient recruitment and helping ensure follow-up data were returned.

References

1 Dale JJ, Callam MJ, Ruckley CV, Harper DR, Berrey PN. Chronic ulcers of the leg: a study of prevalence in a Scottish community. *Health Bull* 1983; **41**: 310–314.
2 Roe B, Cullum N, Hamer C. Patients' perceptions of chronic leg ulceration. In *Leg Ulcers: Nursing Management*, Cullum N, Roe B (eds). Scutari Press: London, 1995; 125–134.
3 Laing W. *Chronic Venous Diseases of the Leg*. Office of Health Economics: London, 1992.
4 Bosanquet N. Costs of venous ulcers: from maintenance therapy to investment programmes. *Phlebology* 1992; **1**(Suppl 1): 44–46.
5 Cullum N, Nelson EA, Flemming K, Sheldon T. Systematic reviews of wound care management: (5) beds; (6) compression; (7) laser therapy, therapeutic ultrasound, electrotherapy and electromagnetic therapy. *Health Technol Assess* 2001; **5**: 1–221.
6 Hampton S. Venous leg ulcers: short-stretch bandage compression therapy. *Br J Nurs* 1997; **6**: 990–992, 994, 996–968.
7 Margolis DJ, Berlin JA, Strom BL. Risk factors associated with the failure of a venous leg ulcer to heal. *Arch Dermatol* 1999; **135**: 920–926.
8 Moffatt CJ, Simon DA, Franks PJ, Connolly M, Fielden S, Groarke L *et al.* Randomised trial comparing two four-layer bandage systems in the management of chronic leg ulceration. *Phlebology* 1999; **14**: 139–142.
9 Moffatt CJ, O'Hare L. Venous leg ulceration: treatment by high compression bandaging. *Ostomy Wound Manage* 1995; **41**: 16–25.
10 Charles H. Short-stretch bandaging in the treatment of venous leg ulcers. *J Wound Care* 1999; **8**: 303–304.
11 Jenkinson C, Layte R, Jenkinson D, Lawrence K, Peterson S, Paice C *et al.* A shorter form health survey: can the SF-12 replicate results from the SF-36 in longitudinal studies? *J Public Health Med* 1997; **19**: 179–186.
12 Hurst NP, Kind P, Ruta D, Hunter M, Stubbings A. Measuring health-related quality of life in rheumatoid arthritis: validity, responsiveness and reliability of the EuroQol (EQ-5D). *Br J Rheumatol* 1997; **36**: 551–559.
13 Hyland ME, Thomson B. Quality of life of leg ulcer patients: questionnaire and preliminary findings. *J Wound Care* 1994; **3**: 294–298.
14 Iglesias CP, Nelson EA, Cullum NA, Torgerson DJ on behalf of the VenUS team. Economic analysis of a randomized trial of two bandages for treating venous leg ulcers. *Br J Surg* 2004; **91**: 1300–1306.
15 Adderley U. Nelson EA. NT Know how: clinical photography. *Nurs Times* 2000; **96**(45): NT Plus 14–15.
16 Altman DG. *Practical Statistics for Medical Research*. Chapman and Hall: London, 1991.
17 Collett D. *Modelling Survival Data in Medical Research. Text in Statistical Science*. Chapman and Hall: London, 1994; 78–85.
18 Grouin JM, Lewis J, Committee for Proprietary Medicinal Products (CPMP). Points to consider on adjustment for baseline covariates. *Stat Med* 2004; **23**: 701–709.
19 Duby T, Cherry G, Hoffman D, Cameron J, Dobloff-Brown D, Ryan T. A randomized trial in the treatment of venous leg ulcers comparing short stretch bandages, four layer bandage system, and a long stretch-paste bandage system. *Wounds – A Compendium of Clinical Research & Practice* 1993; **5**: 276–279.
20 Scriven JM, Taylor LE, Wood AJ, Bell PRF, Naylor AR, London NJM. A prospective randomised trial of four-layer *versus* short stretch compression bandages for the treatment of venous leg ulcers. *Ann R Coll Surg Engl* 1998; **80**: 215–220.
21 Partsch H, Damstra RJ, Tazelaar DJ, Schuller-Petrovic S, Velders AJ, de Rooij MJ *et al.* Multicentre, randomised controlled trial of four-layer bandaging *versus* short-stretch bandaging in the treatment of venous leg ulcers. *Vasa* 2001; **30**: 108–113.
22 Ukat A, Konig M, Vanscheidt W, Munter KC. Short-stretch *versus* multilayer compression for venous leg ulcers: a comparison of healing rates. *J Wound Care* 2003; **12**: 139–143.

Randomized trial

Economic analysis of VenUS I, a randomized trial of two bandages for treating venous leg ulcers

C. P. Iglesias, E. A. Nelson, N. Cullum and D. J. Torgerson on behalf of the VenUS I collaborators

Correspondence to: Professor N. Cullum, Department of Health Sciences, Seebohm Rowntree Building (Area 4), University of York, Heslington, York YO10 5DD, UK (e-mail: nac2@york.ac.uk)

Background: The study investigated the cost-effectiveness of four-layer and short-stretch compression bandages for treating venous leg ulcers.

Methods: Cost-effectiveness and cost-utility analyses were performed using patient-level data collected alongside the VenUS I leg ulcer study. The perspective for the economic analysis was that of the UK National Health Service (NHS) and Personal Social Service. The time horizon for the analysis was 1 year after recruitment. Health benefit was measured as differences in ulcer-free days and quality-adjusted life years (QALYs).

Results: The mean healing time for ulcers treated with four-layer bandages was 10·9 (95 per cent confidence interval (c.i.) −6·8 to 29·1) days less than that for ulcers treated with short-stretch bandages. Mean average difference in QALYs between compression systems was −0·02 (95 per cent c.i. −0·08 to 0·04). The four-layer bandage cost a mean of £227·32 (95 per cent c.i. £16·53 to £448·30) less per patient per year than the short-stretch bandage.

Conclusion: On average, four-layer bandaging was associated with greater health benefits and lower costs than short-stretch bandaging.

Paper accepted 18 May 2004

Published online in Wiley InterScience (www.bjs.co.uk). **DOI:** 10.1002/bjs.4755

Introduction

Leg ulceration remains a common condition, particularly in older women[1]. Studies in people with leg ulceration suggest that quality of life is significantly reduced owing to problems with activity and mobility, pain and lower levels of self-esteem[2].

The different compression bandage systems used to treat leg ulcers have different performance and economic characteristics which make their comparison in a large randomized trial worthwhile. Two high-compression systems are generally employed. The short-stretch system is widely used in continental Europe and the four-layer high-compression bandage system is used in the UK. Although both methods apply high compression, the purchase costs of short-stretch bandages are substantially lower. This study reports an economic analysis of the UK National Health Service (NHS) HTA programme-funded VenUS I, a large multicentre pragmatic randomized trial that compared the clinical and cost effectiveness of four-layer and short-stretch compression bandages for healing venous leg ulcers.

Patients and methods

The perspective for the economic analysis was that of the UK NHS and Personal Social Service[3]. The time horizon for the analysis was 1 year after recruitment.

Study description

Full details of the VenUS I study have been described elsewhere[4]. In brief, 387 patients were recruited to a pragmatic randomized trial from nine geographical areas in Britain. Patients with new, existing or recurrent ulceration of at least 1 week's duration, at least 1 cm in diameter, and with an ankle : brachial pressure index of at least 0·8 were eligible. Patients with diabetes and those whose ulcers had previously failed to improve while using one of the trial bandages were excluded.

The mean age of the patients recruited was 71 (range 23–97) years. The majority of patients in this trial (81·7 per cent; 316 of 387) had a reference ulcer with an area of 10 cm^2 or less.

Data collection

Data on the use of resources, including the reason for the visit and the total number of new bandages used, were collected prospectively by nurses during visits for the primary purpose of leg ulcer care. A questionnaire on health and social care resource use was completed monthly by patients, indicating how many times in the previous month they had used health services (for example, seen a doctor or nurse, received care in hospital, outpatients, community clinic or their own home). As this was a pragmatic trial, the schedule of patient visits was not dictated by trial protocol, except to say that patients should be seen at least once a week.

Cost estimation

Four different types of resource were included in the estimation of costs: number of nurse visits, number of doctor visits, number of hospital visits, and number of bandages used. Information regarding wound cleansers, number and size of primary and secondary dressings, and the use of skin preparations was not included in the economic analysis as these were common to both compression systems, so their effect should cancel out in the incremental analysis of costs[5].

All costs were measured using 2001 prices. Cost and duration of nurse and doctor visits were estimated using the unit cost of community-based healthcare staff[6]. In the base-case analysis it was considered that a nurse visit for leg ulcer care took 22 min in a clinic and 40 min at home[5]. Costs of hospital visits were based on average outpatient costs per day[5]. Acquisition costs of the different bandage systems were taken from the retail prices quoted in the *British National Formulary*[7]. Unit cost estimates used in the economic analysis are described in *Table 1*.

Table 1 Unit costs and sources

	Unit cost*	Reference
Nurse visit		
Home (40 min)	£37	6
Clinic (22 min)	£15	6
Doctor visit		
Home (13·2 min)	£59	6
Clinic (12·6 min)	£26	6
Hospital visit (outpatient) other reasons	£74	6
Four-layer bandage system		
Soffban®	£0.61	7
Softex®	£0.60	7
Velband/Soffband®	£0.72	7
Soffcrepe®	£1.18	7
Setocrepe®	£1.12	7
Crepe	£0.87	7
Litepress®	£3.44	7
Elset®	£2.46	7
Co-plus®	£2.85	7
Coban®	£3.01	7
Four-layer bandage system (kit)	£8.88	Smith & Nephew
Short-stretch bandage system†		
Comprilan®/Rosidal K 12®	£4.08	7
Comprilan®/Rosidal K 10®	£3.43	7
Comprilan®/Rosidal K 12®	£3.08	7
Soffban®	£0.61	7
Softex®	£0.60	7
Velband®	£0.72	7
Other bandages		
Setopress®	£3.29	7
Tensopress®	£3.24	7
Surepress®	£3.18	7
Tubigrip®	£9.00	www.westonshealth.co.uk

*Year of pricing 2001. †Rosidal K® cannot be prescribed by the NHS but could be bought over the counter by nurses or patients, thus it was included in the costing exercise. The official SSB in the VenUS I study was Comprilan®.
Soffban®, Soffcrepe®, Litepress®, Co-Plus®, Tensopress®, Smith & Nephew, Hull, UK; Velband®, Johnson & Johnson, Berkshire, UK; Softexe®, Setocrepe®, Elset®, Setopress®, Tubigrip®, SSL, Cheshire, UK; Coban®, 3M, Leicester, UK; Comprilan®, Beiersdorf, Bucks, UK; Rosidal K®, Vernon-Carus, Preston, UK; Surepress®, Convatec, Uxbridge, UK

Health benefits

Health benefits were measured in two ways. First, Kaplan–Meier estimates of mean time to healing over 12 months in each trial arm were calculated. These in turn were used to estimate the difference in ulcer-free days from the trial data on the basis of intention to treat[4]. Second, quality-adjusted life years (QALYs) were estimated using patients' responses to the EuroQol (EQ) 5D questionnaire. EQ-5D scores for each patient at baseline and every 3 months thereafter during the first year of follow-up were calculated using UK social tariffs[8]. QALYs were calculated for each patient using the area under the curve defined by his or her EQ-5D scores over time. To account for censoring, QALYs were adjusted by the Kaplan–Maier estimates of patients' survival over a year using the method of Lin *et al.*[9].

Economic evaluation analyses

The two compression systems were compared in terms of both the costs and the health benefits associated with their use. The exact form that an economic evaluation takes depends mainly on the way in which health benefits associated with the technology are measured, so that the measurement is both clinically and economically relevant[10]. In this trial two different types of analysis

were conducted: a cost-effectiveness analysis in which health benefits were measured in 'natural' units (ulcer-free days), and a cost-utility analysis in which health benefit measurement was adjusted by changes in quality of life. Taken that the analysis concerned a single year, both costs and benefits were left undiscounted.

Statistical analysis

The economic analysis was conducted using Stata 7. The quantity of resources used per month was summarized using simple descriptive statistics. Given the censored nature of cost data (information regarding a patient's resource use may be truncated at any point in time before the end of the study because of healing of the ulcer, loss to follow-up, or patient's request), the method of Lin *et al.*[9] was used to estimate mean total treatment cost in each trial arm. The event of interest for this analysis was ulcer healing; it was assumed that no further costs occurred after an ulcer had healed.

To account for the skewed nature of cost data, 95 per cent confidence intervals (c.i.) for the adjusted total cost and the average mean cost difference between bandage treatments were estimated using non-parametric bootstrapping techniques (1000 replications). Bias-corrected and accelerated confidence intervals were estimated[11]. Non-parametric bootstrapping techniques were also used to estimate the confidence interval around the estimated mean difference in ulcer-free days, and mean difference in QALYs. Linear multiple regression was used to adjust the differential QALYs for any imbalances in the EQ-5D score at baseline.

Sensitivity analysis

The total costs of compression therapy in the treatment of venous leg ulcers are driven mainly by three items: the number of nurse visits, the setting in which leg ulcer care is delivered, and number of bandages used. The number and setting of nurse visits required by a patient are issues mainly related to the way in which service is provided in the UK, so no sensitivity analysis on these two variables was justified. In urban areas patients with a venous leg ulcer are often treated in a specialized clinic, whereas in rural areas ulcers are treated mainly by domiciliary visits from community nurses. To explore the robustness of the base-case analysis results, the scenario approach to sensitivity analysis was used[12,13]. Three clinically plausible modes of bandage use and supply were explored. In the first scenario, for all the patients allocated to treatment with the short-stretch bandage, the history of resources used per month was analysed individually. Each time bandages were replaced, the quantity of new bandages reported was replaced with the patient's reported minimum number of bandages used over the interval of treatment. The second scenario considered the possibility of acquiring the four-layer compression system as a kit. Although the four-layer kit was not available on prescription in the UK during the trial, leg ulcer clinics commonly purchase kits and supply them from stocks. In the third scenario, the two options described above were considered simultaneously.

Cost-effectiveness planes were used to explore changes in the distribution of the incremental costs and health benefits associated with the different sensitivity analyses. The level of uncertainty associated with the cost-effectiveness of the short-stretch bandage compared with the four-layer bandage was explored using cost-effectiveness acceptability curves. Acceptability curves represent the probability of an alternative being cost effective for a range of willingness to pay values for an extra unit of health benefit associated with the alternative[14].

Results

Information from 387 patients was included in the analysis, 195 allocated to the four-layer bandages (4LB) group and 192 allocated to the short-stretch bandages (SSB) group.

Effectiveness

In a Kaplan–Meier survival analysis of cumulative healing times there was a trend towards increased healing in the 4LB group, although this was not statistically significant (log rank $\chi^2 = 2{\cdot}46$, $P = 0{\cdot}117$). Analysis using a Cox proportional hazards model to adjust for the effects of other variables that may influence healing (treatment centre, baseline ulcer area, duration, episodes, ankle mobility, weight) suggested a statistically significant treatment effect in favour of four-layer bandaging. The hazard ratio for healing, 0·72 (95 per cent c.i. 0·57 to 0·91), suggested that individuals in the SSB group had a significantly lower probability of healing than those in the 4LB group.

Costs

Descriptive statistics of monthly volume of resources used in both bandage groups are shown in *Table 2*. Between one and two extra nurse visits per month were required by patients in the SSB group. No marked differences in monthly volume of other resource items were observed between groups.

A summary of the adjusted total mean annual costs estimated for both treatments is presented in *Table 3*.

Table 2 Monthly resources used per trial arm during the first year of follow-up

Resource	4LB	SSB
Nurse visits		
Mean(s.d.)	5·14(3·26)	6·03(4·09)
Median (range)	4 (0–21)	5 (1–29)
Doctor visits		
Mean(s.d.)	0·23(2·17)	0·34(1·44)
Median (range)	0 (0–62)	0 (0–14)
Hospital visits		
Mean(s.d.)	0·40(1·04)	0·33(1·10)
Median (range)	0 (0–7)	0 (0–15)
Wool bandages		
Mean(s.d.)	4·93(3·55)	5·87(5·61)
Median (range)	4 (0–22)	4 (0–54)
Crepe bandages		
Mean(s.d.)	3·34(2·83)	0·79(2·30)
Median (range)	3 (0–18)	0 (0–20)
Elset bandages		
Mean(s.d.)	3·57(2·94)	0·59(1·86)
Median (range)	4 (0–20)	0 (0–17)
Coban bandages		
Mean(s.d.)	4·08(3·02)	0·83(2·22)
Median (range)	4 (0–20)	0 (0–16)
Comprilan 12 bandages		
Mean(s.d.)	0·00(0·14)	0·48(2·31)
Median (range)	0 (0–3)	0 (0–25)
Comprilan 10 bandages		
Mean(s.d.)	0·08(0·92)	2·71(4·88)
Median (range)	0 (0–16)	0 (0–52)
Comprilan 8 bandages		
Mean(s.d.)	0·01(0·19)	0·60(1·86)
Median (range)	0 (0–4)	0 (0–18)
Other bandages		
Mean(s.d.)	0·03(0·30)	0·03(0·41)
Median (range)	0 (0–5)	0 (0–9)

4LB, four-layer bandages; SSB, short-stretch bandages.

These estimates showed a statistically significant difference in favour of the four-layer bandage. Treatment with four-layer bandages cost a mean £227·32 (95 per cent c.i. £16·53 to £448·30) less per patient per year. There was no statistically significant difference in Kaplan–Meier estimates of the mean time to healing over a year; ulcers in the 4LB group healed a mean of 10·9 (95 per cent c.i. – 6·76 to 29·06) days before those in the SSB group (*Table 3*).

Mean average adjusted QALYs gained per bandage group are shown in *Table 3*. Having adjusted for baseline utility scores and censoring, there was no statistically significant difference in the QALYs after the first year. Individuals in the 4LB group had, on average, a better quality of life than those in the SSB group; the annual difference in QALYs was – 0·02 (95 per cent c.i. – 0·08 to 0·04).

The base-case analysis showed that four-layer bandaging was the superior strategy, and was associated on average with a greater health benefit and lower costs than the short-stretch bandage; an incremental analysis was therefore not justified.

Sensitivity analysis

Point estimates of the difference in costs between four-layer and short-stretch bandages for the three scenarios are shown in *Table 4*. In all three scenarios the difference in total treatment cost over a year was smaller than that estimated in the base-case scenario (£227·32). In none of the three scenarios considered was the difference in costs statistically significant at the 5 per cent level (*Table 4*).

Table 3 Mean adjusted annual costs and health benefits

	4LB	SSB	Difference
Costs (£)	1298·41 (1187·83, 1471·90)	1525·73 (1373·92, 1716·66)	227·32 (16·53, 448·30)
Time to heal (days)	96·7 (85·3, 111·6)	107·6 (95·9, 122·2)	10·9 (– 6·8, 29·1)
QALYs	0·69 (0·66, 0·74)	0·67 (0·63, 0·72)	– 0·02 (– 0·08, 0·04)

Values in parentheses are 95 per cent bias-corrected and accelerated confidence intervals. 4LB, four-layer bandages; SSB, short-stretch bandages; QALY, quality-adjusted life year.

Table 4 Mean adjusted annual costs for different scenarios

	4LB	SSB	Difference
First scenario (£)	1298·41 (1187·83, 1471·90)	1486·47 (1339·77, 1668·03)	188·06 (– 15·36, 410·00)
Second scenario (£)	1385·92 (1270·11, 1565·79)	1525·73 (1373·91, 1716·66)	139·81 (– 62·11, 369·09)
Third scenario (£)	1385·92 (1270·11, 1565·79)	1486·47 (1339·77, 1668·03)	100·55 (– 105·13, 321·62)

Values in parentheses are 95 per cent bias-corrected and accelerated confidence intervals. 4LB, four-layer bandages; SSB, short-stretch bandages.

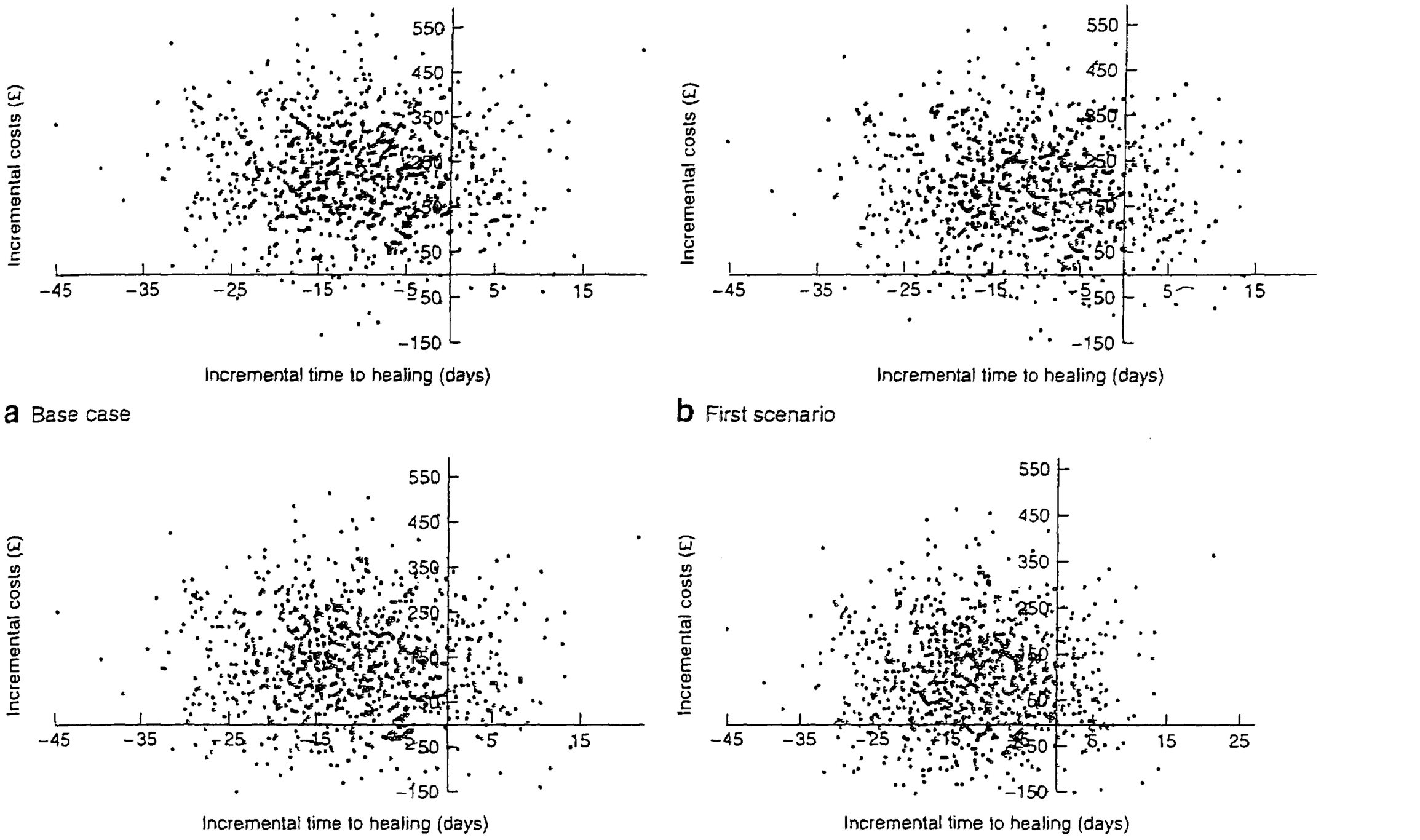


Fig. 1 Cost-effectiveness plane (sensitivity analysis)

The robustness of the results was tested using a scenario approach to sensitivity analysis, which indicated that even if the quantity of short-stretch bandages used in the treatment of a patient was lower than that reported in this trial, and the acquisition price of the four-layer bandage was at its maximum level, on average the four-layer bandage was still the dominant alternative. Cost-effectiveness planes for the three sensitivity analysis scenarios are compared with the base-case analysis in *Fig. 1*.

The main effect of considering feasible variations in the number of bandages used and the unit costs of the compression systems was to reduce the mean difference in costs between the two treatments, shifting the distribution of point estimates towards the third quadrant, where short-stretch compared with four-layer bandaging was still associated with less health benefit but now also fewer costs. However, in all three scenarios the vast majority of point estimates fell in the upper left quadrant, suggesting that even under these conditions four-layer bandaging was a dominant strategy. The cost-effectiveness acceptability curve for the short-stretch bandage suggested that, even for considerably high willingness to pay values for an ulcer-free day, the short-stretch bandage was associated with only a 20 per cent probability of being cost effective.

Discussion

Venous leg ulcer is a costly chronic condition. Management is currently provided within specialized clinics and in the home by community nurses. Given the large amount of human and medical resources that such services demand, it is essential to ensure that provision of healthcare is clinically and economically efficient. Compression therapy is considered to be the key component in the effective treatment of venous ulcers. The cost-effectiveness of alternative methods of service provision has been investigated previously[15]. This analysis investigated the cost-effectiveness of short-stretch compared with four-layer bandaging, using two different methods of measuring the health benefit associated with these two compression systems: natural units (cost-effectiveness) and quality of life (cost-utility analysis).

Both economic analyses showed that short-stretch bandaging was inferior to four-layer bandaging; on average

short-stretch bandaging was associated with smaller health benefits and greater costs. Patients in the 4LB group healed a mean of 10·9 days before those in the SSB group and the mean average difference in QALYs between compression systems was 0·02. The mean difference in total cost between compression systems was £227·32 per patient per year in favour of four-layer bandaging. Although some of these differences were not statistically significant at the 5 per cent level, it has been argued, from an economic perspective, that decisions should be made on the basis of the expected mean difference in health benefits and costs between the alternative programmes under evaluation, irrespective of their statistical significance[16].

On average, patients in the SSB group required one more visit per month than patients in the 4LB group. Across all centres most care was delivered via domiciliary visits. This, combined with the substantial difference between the cost of clinic (£15) and home (£37) care, meant that even slight imbalances in the number of visits between the bandage groups resulted in a substantial increase in the total costs of the treatment. According to the information from the manufacturers, both compression systems were able to remain in place for 1 week, so there was no reason to believe that there would be an imbalance in the number of visits. Consequently no detailed information regarding the reasons for the increased frequency of nurse visits was collected. In future studies, the factors that determine the frequency of nurse contacts with patients should be investigated.

The number of short-stretch bandages used per patient per month was substantially higher than that recommended by the manufacturers; this recommendation was based on the effect of laundering on bandage performance. According to nurse reports, the higher usage of these bandages was mainly explained by the characteristics of the study population; elderly patients often found it difficult to wash the bandages, and nurses preferred not to reuse the bandages from patients with sloughy, exuding ulcers. This showed the importance of undertaking an economic analysis alongside a clinical trial, as these economically significant data were only revealed within the context of a pragmatic trial.

The centres that participated in this trial had a range of methods of delivering venous leg ulcer care in the community. Specialist leg ulcer services, integrated vascular services and community nurse-led services were all represented, and so the findings of this study can be applied across the UK. However, when attempting to extrapolate from these results to individual leg ulcer services there are important considerations to be made, particularly with regard to existing skills and competencies in bandage application as the relative clinical effectiveness of the compression systems is likely to be heavily influenced by these factors.

In this analysis the relative cost-effectiveness of two compressions systems to resolve an acute episode of ulceration was explored over a 1-year interval. Given the chronic nature of venous leg ulcers, it can be argued that to explore the cost-effectiveness of alternative methods of managing venous leg ulcers one should take into account not only the resolution of isolated episodes of ulceration, but also consider the health benefits and costs associated with preventive measures, and the frequency of new ulceration episodes. Such an analysis should be the subject of future modelling exercises, rather than the result of an evaluation conducted alongside a randomized study.

The present analysis supported the use of four-layer bandages in preference to short-stretch bandages, on the basis of clinical and economic superiority. However, for centres in which short-stretch bandaging is currently used, a careful audit of the frequency of district nurse visits and healing rates is recommended. If the size of a leg ulcer is reducing, but domiciliary visits for short-stretch rebandaging are required more than once per week, then the care is not cost effective. However, if the ulcer size is reducing, and patients and/or their carers are able to launder and reapply the bandage, then the short-stretch bandage is likely to become cost effective. Short-stretch bandaging remains a reasonable alternative for patients who prefer it or will not tolerate four-layer bandages.

Acknowledgements

The authors thank Andrea Manca for valuable comments on the economic analysis. Other acknowledgements, and a list of all collaborators and funding arrangements, can be found in the clinical results article (*Br J Surg* 2004; **91**: 1292–1299).

References

1 Callam MJ, Harper DR, Dale JJ, Ruckley CV. Chronic ulcer of the leg: clinical history. *BMJ* 1987; **294**: 1389–1391.
2 Roe B, Cullum N, Hamer C. Patients' perceptions of chronic leg ulceration. In *Leg Ulcers: Nursing Management: a Research-Based Guide*, Cullum NA, Roe B (eds). Scutari Press: Harrow, 1995; 125–134.
3 National Institute for Clinical Excellence. *Technical Guidance for Manufacturers and Sponsors on Making a Submission to a Technology Appraisal*. National Institute for Clinical Excellence: London, 2001.
4 Nelson EA, Iglesias C, Cullum NA, Torgerson DJ on behalf of the VenUS I collaborators. Randomized clinical trial of

four-layer and short-stretch compression bandages for venous leg ulcers (VenUS I). *Br J Surg* 2004; 91: 1292–1299.
5 Drummond MF, O'Brien BJ, Stoddart GL, Torrance GW. *Methods for the Economic Evaluation of Health Care Programmes* (2nd edn). Oxford University Press: Oxford, 1997.
6 Netten A, Curtis L. *Unit Cost of Health and Social Care*. Personal and Social Services Research Unit, University of Kent: Canterbury, 2001.
7 British Medical Association/Royal Pharmaceutical Society of Great Britain. *British National Formulary 42: September 2001*. British Medical Association/Royal Pharmaceutical Society of Great Britain: London, 2001.
8 Dolan P, Gudex C, Kind P, Williams A. *A Social Tariff for EuroQol: Results from a UK General Population Survey*. University of York Centre for Health Economics Discussion Paper Series (no. 138). Centre for Health Economics: York, 1995.
9 Lin DY, Feuer EJ, Etzioni R, Wax Y. Estimating medical costs from incomplete follow-up data. *Biometrics* 1997; 53: 419–434.
10 Iglesias C, Torgerson DJ. Health economics. In *Medicine in Society: Behavioural Sciences for Medical Students*, Dowrick C (ed.). Arnold: London, 2001; 171–180.
11 Efron B, Tibshirani RJ. *An Introduction to the Bootstrap*. Chapman & Hall: New York, 1993.
12 Briggs A, Sculpher M, Buxton M. Uncertainty in the economic evaluation of health care technologies: the role of sensitivity analysis. *Health Econ* 1994; 3: 95–104.
13 Briggs AH, Gray AM. Handling uncertainty when performing economic evaluation of healthcare interventions. *Health Technol Assess* 1999; 3: 1–134.
14 van Hout VA, Al MJ, Gordon GS, Rutten FF. Cost, effects and C/E ratios alongside a clinical trial. *Health Econ* 1994; 3: 309–319.
15 Morrell CJ, Walters SJ, Dixon S, Collins KA, Brereton LM, Peters J *et al*. Cost effectiveness of community leg ulcer clinics: randomised controlled trial. *BMJ* 1998; 316: 1487–1491.
16 Claxton K. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. *J Health Econ* 1999; 18: 341–364.

ORIGINAL RESEARCH ARTICLE

Pharmacoeconomics 2006; 24 (5): 465-478
1170-7690/06/0005-0465/$39.95/0

Comprehensive Decision-Analytic Model and Bayesian Value-of-Information Analysis

Pentoxifylline in the Treatment of Chronic Venous Leg Ulcers

Cynthia P. Iglesias[1,2] and *Karl Claxton*[2,3]

1 Department of Health Sciences, University of York, York, UK
2 Centre for Health Economics, University of York, York, UK
3 Department of Economics and Related Studies, University of York, York, UK

Abstract

Objective: To conduct a Bayesian value-of-information analysis of the cost effectiveness of pentoxifylline (vs placebo) as an adjunct to compression for venous leg ulcers.

Methods: A probabilistic Markov model was developed to estimate mean clinical benefits and costs associated with oral pentoxifylline (400mg three times daily) and placebo. Clinical data were obtained from a systematic review and synthesised using Bayesian methods. The decision uncertainty associated with the adoption of pentoxifylline as well as the maximum value associated with further research were estimated before and after the completion of the largest 'definitive' treatment trial. Resource use was obtained from a UK national audit and unit costs applied (£, 2004 values).

Results: The prior and posterior analyses suggest that pentoxifylline is a dominant therapy versus placebo. In the prior analysis, patients in the pentoxifylline group healed an average of 8.28 weeks quicker than patients in the placebo group (95% credibility interval [CI] 1.89, 14.56), had a 0.02 gain in QALYs (95% CI −0.12, 0.17) and an average reduction in cost of £153.4 (95% CI −53.11, 354.9).

Estimates of the uncertainty surrounding the cost effectiveness of pentoxifylline and the value of perfect information in both analyses did not suggest further research was justified. In the prior analysis, for willingness-to-pay values of £0, £100 and £500 per QALY gained, the estimated values of perfect information were £128 200, £127 100 and £126 700, respectively.

Incorporation of the information from the largest randomised controlled trial on pentoxifylline did improve the estimate of the clinical effect associated with this drug; however, the variation was not large enough to reverse either the decision regarding the dominance of pentoxifylline or the maximum value associated with further research.

Conclusion: Bayesian value-of-information analysis represents a valuable tool for healthcare decision making. Had the results from this analysis been available before the largest trial was funded, a more efficient allocation of research and development resources could have been made.

Randomised controlled trials (RCTs) are considered the 'gold standard' for collecting information about the clinical and cost effectiveness of healthcare interventions. However, large confirmatory RCTs are often expensive and difficult to undertake, and can take considerable time. In therapeutic areas where a definitive treatment trial is not available, systematic reviews and meta-analyses of existing evidence from (often insufficiently powered) trials are commonly used as an interim solution to the lack of evidence. However, it is widely believed that the results of a systematic review and meta-analysis ought to be confirmed subsequently in a large rigorously designed RCT.

The replication of RCTs has frequently been justified by the need to demonstrate the clinical effectiveness of a treatment at conventional levels of statistical significance (i.e. $p < 0.05$). However, it has been argued that, in the context of healthcare decision making, traditional rules of statistical inference, from either a Bayesian or frequentist viewpoint, are not consistent with the objective of a publicly funded healthcare system that operates under a fixed budget.[1] If the objective is to maximise health benefit within a given budget, decision makers in the health sector should base their decisions regarding the adoption of new health technologies on the mean 'net benefit', which is defined as a linear function of the expected clinical benefits and costs associated with the technology.[1,2] By rejecting a new health technology associated with a positive mean 'net benefit' merely on the grounds of a lack of statistical significance, healthcare decision makers will be automatically surrendering the potential mean clinical benefit associated with the new technology.[1]

Research funding bodies (such as the UK NHS Health Technology Assessment Programme and the Medical Research Council) routinely consider the scientific excellence of a research proposal, and often require evidence of the degree of uncertainty regarding clinical and cost effectiveness of health technologies through systematic reviews. An estimate of the actual value associated with the reduction of this uncertainty is not a formal requirement. This current procedure ignores the fixed cost of research and places an unlimited value on its associated benefits, that is, the reduction in decision uncertainty. If the maximum value associated with the reduction in decision uncertainty is exceeded by the fixed cost of research, the acquisition of further information is unlikely to be an efficient use of already scarce research and development resources.[1] Furthermore, undertaking research in one therapeutic area has a direct opportunity cost because it deprives another research area of vital evaluative resources.

It has been argued that the question of whether additional clinical trial research is needed ought to be informed by an assessment of the relative costs and benefits of undertaking such a trial.[3] A formal framework, based on decision-analytic models and Bayesian value-of-information analysis, to estimate the maximum value associated with additional research has been proposed.[4,5] This approach can inform two key stages in the process of deciding whether to recommend the adoption of a health technology: (i) whether a health intervention should be considered clinically and cost effective given existing information and; (ii) whether a request for further information is justified. In other words, this framework enables healthcare researchers to address decisions regarding an efficient provision of healthcare services based on available evidence, and helps to identify those areas where the acquisition of further information, through primary research, is more likely to reduce the uncertainty surrounding the decision about whether or not to adopt a new healthcare technology.

In this paper we demonstrate the application of the above methods by conducting a Bayesian value-of-information analysis to evaluate the cost effectiveness of oral pentoxifylline (400mg three times daily) compared with placebo as an adjunct to compression therapy in the treatment of chronic venous ulcers. The objectives of this analysis were as follows.

1. To explore the decision uncertainty regarding the cost effectiveness of oral pentoxifylline (400mg three times daily), based on the available evidence before the publication of the largest study conducted in this area; we shall refer to this as the 'prior analysis'.
2. To evaluate the efficiency of the decision made to mount the largest RCT on oral pentoxifylline (400mg three times daily).

3. To determine the contribution of the findings from the latter trial to reduce the decision uncertainty regarding the cost effectiveness of oral pentoxifylline (400mg three times daily); we shall refer to this as the 'posterior analysis'.
4. To provide information regarding future research on pentoxifylline.

Case Study

Venous leg ulcers are a chronic and recurrent condition affecting approximately 1.5 per 1000 of the UK population aged ≥18 years or 62 000 people with a history of leg ulceration at any time.[6] Venous leg ulcers are managed primarily by nurses in hospitals, leg ulcer clinics and in the community. In 1992 the total annual cost for the UK NHS for venous ulcers was estimated at around £400 million.[7]

Current practice in the treatment of venous leg ulcers involves the application of topical agents (e.g. wound dressings, hydro gels, etc.) and compression therapy (e.g. two-, three- and four-layer bandages, Unna's boot or stockings). Whilst the aetiology of the disease has not been explained fully, most leg ulcers are associated with venous disease. Deep vein thrombosis is thought to cause damage to the deep veins, leading to venous insufficiency and ulceration. Whilst no pharmacological treatment is indicated for venous leg ulcers, it has been suggested that improving venous deficiencies may accelerate ulcer healing.[8] Pentoxifylline, a drug that helps blood flow, has been postulated as a potential pharmacological agent to speed up the healing process of venous leg ulcers. A systematic review of the evidence on pentoxifylline for the treatment of venous leg ulcers concluded that this drug provides additional benefits to compression and is possibly effective for patients not receiving compression.[9]

The largest RCT conducted on pentoxifylline to date was published in 1999 by Dale et al.[10] Prior to the dissemination of the results from this study, three RCTs provided evidence on the effectiveness of oral pentoxifylline (400mg three times daily), as an adjunct to compression, relative to placebo in the treatment of chronic venous leg ulcers (table I). These three studies randomised 232 participants in total.[11-13] A statistically significant positive effect was only reported in the study by Colgan et al.[12] However, using a Bayesian random baseline fixed-effect meta-analysis to estimate a pooled odds ratio of healing at 24 weeks with the data from these three studies, a statistically significant benefit for pentoxifylline versus placebo was calculated at 2.12 (95% credibility interval [CI] 1.29, 3.34). However, the small sample sizes, unclear descriptions of randomisation method and short follow-up periods reported raised concerns regarding the validity and reliability of the findings from these studies.

The study by Dale et al.,[10] which addressed many of the methodological shortcomings associated with previous studies, found a positive but non-statistically significant effect for pentoxifylline (vs placebo) on ulcer healing rates (table I). When the results from Dale et al.[10] were included in the meta-analysis described above, the estimate of the pooled odds ratio of healing at 24 weeks changed to 1.93 (95% CI 1.41, 2.61). Whilst the results from the trials by Falanga et al.[13] and Dale et al.[10] were published in the same year (1999), we shall assume for illustrative purposes that the findings from

Table I. Main characteristics of randomised placebo-controlled studies on pentoxifylline (Pent) 1200mg daily in the presence of compression for the treatment of venous leg ulcers

Study[a]	Intervention group (n)	Compression type (n[b])	Mean age (y)	Follow-up (wk)	Ulcer healing	
					Pent	placebo
Falanga et al.[13]	Pent (83)	Unna's boot (45)	58	24	62/83	28/45
Colgan et al.[12]	Pent (38)	Two-layer system (42)	71	24	23/38	12/42
Schurmann and Eberhardt[11]	Pent (12)	Unspecified (12)	65	8	2/12	3/12
Dale et al.[10]	Pent (101)	Single- or four-layer system (99)	69	24	65/101	52/99

a The prior analysis included data from Falanga et al.,[13] Colgan et al.[12] and Schurmann and Eberhardt.[11] The posterior analysis included data from Dale et al.[10] in addition to the other three studies.

b n refers to number of patients treated with compression alone in the placebo group; all patients treated with pentoxifylline also received compression measures.

Falanga et al.[13] were available at the time that the Dale et al.[10] study was submitted for funding.

Methods

A probabilistic Markov model was used to estimate expected mean clinical benefits and costs associated with two alternative treatments: (i) oral pentoxifylline 400mg three times daily in addition to 'usual wound care' (compression therapy and wound dressings); or (ii) placebo in addition to 'usual wound care'. Venous leg ulceration prognosis was represented by three mutually exclusive health states: healed, non-healed and death (figure 1). Patients were allowed to transit between health states at weekly periods. Leg ulceration recurrence was represented by allowing patients to transit between the healed and unhealed states in both directions. Total follow-up for the base-case analysis was 52 weeks (it was considered unlikely that pentoxifylline would be used for >1 year). Given that the time horizon for the analysis was 1 year, estimated total costs and benefits were not discounted. Table II describes the sources and characterisation of all stochastic and deterministic model parameters.

To simultaneously address the four stages of decision-analytic modelling[5] (systematic review of relevant data incorporating a meta-analysis, estimation of model inputs, sensitivity analysis for data and model specifications and evaluation of the model), a comprehensive decision-analytic model was constructed and evaluated using Markov Chain Monte Carlo simulation implemented in the specialist software WinBUGS.[20]

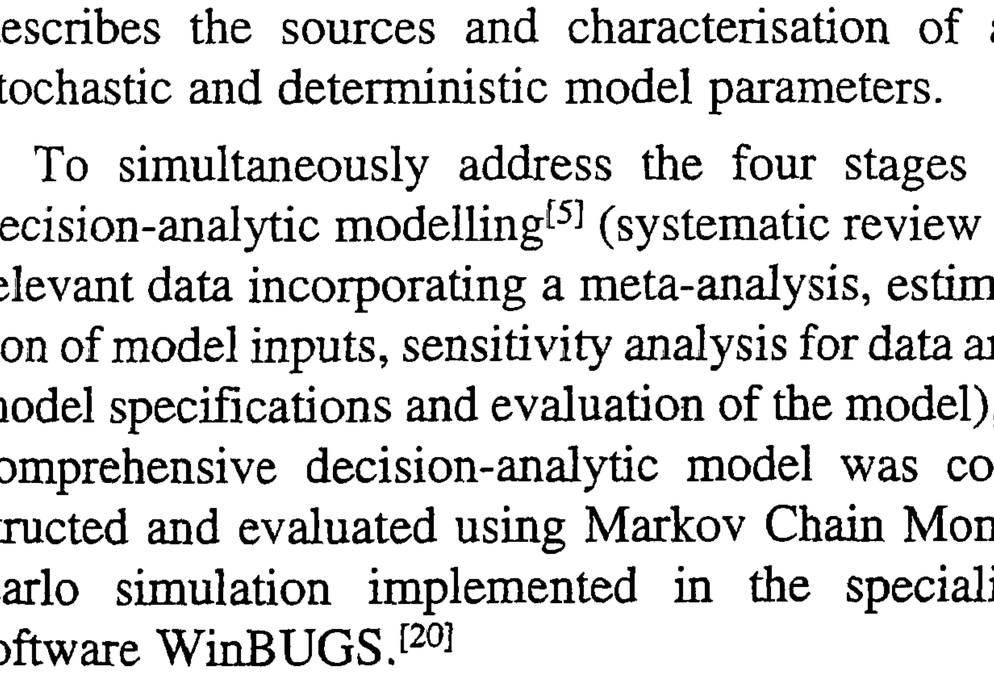


Fig. 1. Stochastic Markov model: graphical description of health states for venous leg ulceration.

Systematic Review and Meta-Analysis

The search strategy described in the Cochrane systematic review[9] of pentoxifylline for venous leg ulcers was used to identify all existing RCTs published before January 2000 comparing oral pentoxifylline with placebo in addition to compression. The Dale et al.[10] study was published in October 1999 and was therefore included in the review.

Four RCTs were identified (see table I).[10-13] In the prior analysis, the results from Schurmann and Eberhardt,[11] Falanga et al.[13] and Colgan et al.[12] were included in a meta-analysis performed on the odds-ratio scale. Despite some expected variability between studies, there was no reason to believe that the effectiveness of pentoxifylline should differ between trials. The potential heterogeneity was reflected into the meta-analysis by using a random-effect model to estimate the natural logarithm of the pooled odds of healing in the placebo group and a fixed-effect model to estimate the natural logarithm of the pooled odds of healing with pentoxifylline. A posterior analysis that included the data from Dale et al.[10] was then conducted to update the pooled odds ratio of healing.

The structure of the Bayesian random baseline fixed-effect meta-analysis used is outlined in equation 1:

Placebo	Pentoxifylline
$r_i^c \sim binomial(n_i^c, p_i^c)$	$r_i^t \sim binomial(n_i^t, p_i^t)$
$\text{logit}(p_i^c) = \theta_i^c$	$\text{logit}(p_i^t) = \theta_i^c + \xi$
$\theta_i^c \sim normal(\mu, \tau^2)$	$or^t = \exp(\mu + \xi)$
$or^c = \exp(\mu)$	$\xi \sim normal(0.0, 1.0E-6)$
$\mu \sim normal(0.0, 1.0E-6)$	
$\tau^2 = \frac{1}{\sigma^2}$	
$\sigma \sim normal(0.2, 400)$	

(Eq. 1)

Where r_i^j is the number of healed individuals of n_i^j in the ith study in group j = placebo (c) and pentoxifylline (t); p_i^j is the estimated healing rates in the ith study and jth groups; or^j are the odds of healing in the jth groups; and σ^2 is the between-study baseline variability or heterogeneity. This parameter estimates the variation in the healing rates observed in the placebo groups that may exist be-

Table II. Characterisation of model parameters (costs are in 2004 values)

Parameter	Prior distribution	Source
Stochastic variables[a]		
Odds of healing	$\mu \sim normal(0.0, 1.0E-6)$ $\delta \sim normal(0.2, 400)$	Schurmann and Eberhardt (1986)[11] Falanga et al. (1999)[13] Colgan et al. (1990)[12] Dale et al. (1999)[10]
Probability of death	$\mu \sim normal(0.0, 1.0E-6)$	Actuary's Government Statistics Interim life tables (2003)[14]
Probability of recurrence	$\mu \sim normal(0.0, 1.0E-6)$	Franks et al. (1995)[15]
Probability of using alternative types of dressings; six alternatives[b]	$d \sim multinomial(p,n)$ $p \sim dirichlet(1,1,1,1,1,1)$ n = total number of responders	RCN National Audit (2000)[16] [c]
Probability of using alternative types of bandages; nine alternatives[d]	$b \sim multinomial(p,n)$ $p \sim dirichlet(1,1,1,1,1,1,1,1,1)$ n = total number of responders	RCN National Audit (2000)[16]
Number of visits; three frequencies[e]	$v \sim multinomial(p,n)$ $p \sim dirichlet(1,1,1,)$ n = total number of responders	RCN National Audit (2000)[16]
Nurse visit duration; three categories[f]	$t \sim multinomial(p,n)$ $p \sim dirichlet(1,1,1,)$ n = total number of responders	RCN National Audit (2000)[16]
Nursing grade; five different grades[g]	$g \sim multinomial(p,n)$ $p \sim dirichlet(1,1,1,1,1)$ n = total number of responders	RCN National Audit (2000)[16]
Visit setting; two potential settings[h]	$s \sim multinomial(p,n)$ $p \sim dirichlet(1,1)$ n = total number of responders	RCN National Audit (2000)[16]
Deterministic variables		
Bandage mean cost (£)[d]	3.07, 2.65, 3.04, 3.28, 6.16, 3.51, 7.60, 21.25, 6.32	BNF September (2004)[17]
Dressing mean cost[b]	2.73, 5.69, 3.31, 2.29, 0.3, 2.86	BNF September (2004)[17]
Nurse mean salary[g]	17 945, 20 122.50, 23 200, 26 447.50, 29 305	Booth (2003)[18]
Discount rate	6% per annum	UK Treasury[19]
Research life cycle	10 years	Expert opinion
Venous ulcer incidence	1.5 per 1000 population aged ≥18y	Callam et al. (1985)[6]
Population projections	Projections for years 2004–13 ranged from 59.5 to 60.7 million	Population projections by the UK Government Actuary (2003-based)[14]

a All normal distributions are characterised in terms of their mean and precision. Multinomial distributions are characterised in terms of sample size and probability parameters re-scaled so that their sum equals 1.

b Dressing type: NA, hydrocolloid, foam, alginate, tulle grass, other.

c RCN, *unpublished* data.

d Bandage type: Setopress®, Surepress®, Tensopress®, Comprilan®, Ultra Four, Rosidal K®, four-layer, hosiery class 2/3, other.

e Frequency of visits: one, two or three visits per week.

f Nurse visit duration: (a) ≤20 min, (b) 30 min, (c) ≥40 min.

g Nurse grades: D, E, F, G, H.

h Visit setting: clinic or community.

BNF = *British National Formulary*; **NA** = knitted viscose primary dressing; **RCN** = Royal College of Nursing.

tween studies. It has been argued that a small number of studies do not provide sufficient evidence to accurately estimate the between-study standard deviation; in these circumstances substantial judgement is required.[21,22] Following a previously published example, we assumed that variations in the system of compression used produced considerable heterogeneity in the baseline healing rates between

studies.[23] Thus σ^2 was assigned a prior normal distribution with mean 0.2 and standard deviation 0.05. This corresponds to expecting ±50% variability in true odds ratios between studies, with 95% uncertainty limits of 20–80%.[22]

Estimation of Model Input Parameters

Transition Probabilities

The probabilities of healing with pentoxifylline and placebo were estimated using equation 2:

$$ph^c = \frac{or^c}{(1+or^c)} \qquad ph^t = \frac{or^t}{(1+or^t)}$$

$$phw^c = 1-(1+ph^c)^{1/24} \qquad phw^t = 1-(1+ph^t)^{1/24} \qquad \text{(Eq. 2)}$$

Where *phj* represents the 24-week cumulative probability of healing with j = placebo (c) and pentoxifylline (t) and *orj* are the odds of healing in the *j*th group estimated from the meta-analysis. Weekly transition probabilities *phwj* were estimated from the cumulative probabilities of healing assuming that the true transition probabilities remained constant over the 24-week period. No residual effect of pentoxifylline was assumed (i.e. once individuals in the intervention group had experienced a healing episode, their probabilities to transit between health states, e.g. moving from unhealed to healed etc, became those of the placebo group).

Probability of Recurrence

Given the chronic nature of venous leg ulcers, the probability of an ulcer re-occurring had to be incorporated into the model. A weekly probability of recurrence was estimated from an 18-month (78-week) rate of recurrence reported in a previously published study.[15] The odds of recurrence were estimated using a fixed-effect model; in turn this was used to estimate the weekly probability of recurrence using equation 3:

$$pr = \frac{or}{(1+or)}$$

$$prw = 1-(1+pr)^{1/78} \qquad \text{(Eq. 3)}$$

Where *or* is the odds of recurrence estimated in a fixed-effect meta-analysis, *pr* is the 18-month probability of recurrence and *prw* is the weekly probability of recurrence.

Probability of Death

Whilst leg ulcers are not a life-threatening condition, they mainly occur in the elderly. Therefore annual mortality rates for the British general population aged 66 years were used to estimate the weekly probability of dying, following the same methodology used to estimate probabilities of recurrence.[14] The age was set at 66 years to reflect the average age over the four studies of pentoxifylline.

Health Benefits

Effectiveness was measured in terms of both QALYs and ulcer-free weeks. The utility weights associated with the healed and unhealed states were assigned beta distributions, the parameters of which were based on data from a previously published RCT and estimated using the methods of moments.[24,25]

Resource Use

To estimate the likely volume of resources used in the treatment of venous leg ulcers we used data from a UK national audit (Royal College of Nursing, unpublished data).[16] Total number of resources used per week in the placebo and pentoxifylline groups were estimated in terms of the number and type of bandages used and dressings applied, number and duration of nurse visits to patients in the community and in a clinical setting and nurse grading in the first and subsequent visits.

The audit provided detailed information about the frequency with which these resources were usually required in the treatment of venous leg ulcers. In turn, this allowed us to stochastically characterise relevant costing items; for example, the mean number of nurse visits per week was estimated as a weighted average of the three potential frequencies of visits considered in the national audit (i.e. one, two or three visits per week). The weights were assumed to follow a multinomial distribution. An uninformative prior conjugate Dirichlet distribution was assigned to the parameter of the multinomial (table II). A similar procedure was followed to estimate type of bandage and dressing used, treatment setting and duration and nursing grading for patients' first assessments and subsequent visits.

Nursing salaries were characterised as uniform probability distributions defined in the range of the minimum and maximum salary points for each

grade. Annual salaries were obtained from the review body for nursing staff, midwives, health visitors and professions allied to medicine.[18]

The frequency of dressing and bandage changes was determined by the number of visits per week. Total weekly costs were calculated by applying average unit prices to total volume of resources utilised in each strategy. Unit prices were obtained from the British National Formulary.[17] The year of pricing was 2004.

Sensitivity Analysis

The sensitivity of the pooled estimate of clinical effectiveness to variations in total length of follow-up between studies was explored in a sensitivity analysis, where weekly transition probabilities of healing were estimated from the proportion of patients healed at 8 weeks rather than the 24-week data.

Sensitivity analyses were also conducted on prior distributions assigned to the following parameters: mean pooled odd ratios of healing, and heterogeneity between studies.

Model Evaluation

Markov Model

To evaluate the Markov model, a cohort of 1000 individuals aged 66 years was entered into the model. At each cycle, the total number of individuals in each state was multiplied by the state's associated utility weights and costs, and then summed across all states and cycles to estimate mean QALYs and mean total cost in the placebo and pentoxifylline groups. This model was evaluated simultaneously with the meta-analysis and the decision tree.

According to preliminary runs, it was decided to use a burn-in of 40 000 iterations in order to reach convergence, and base inferences on a sample of a further 10 000 iterations. Model convergence was confirmed using two sets of rather dispersed starting values and by visual examination of the plot of the Gelman Rubin statistic in WinBUGS.[24] The two chains provided similar results, suggesting convergence had been achieved.

Cost Effectiveness

A cost-utility (base-case) and a cost-effectiveness analyses were conducted. Cost effectiveness was determined using the net benefit approach, according to which the optimal alternative is the one that yields the highest expected 'net benefit'. Net benefits in monetary terms were estimated in equation 4:[2]

$$NB(t,\theta) = \lambda * B(t,\theta) - C(t,\theta) > 0 \quad \text{(Eq. 4)}$$

where NB is the net benefit in monetary terms under treatment t, λ is the decision maker's willingness to pay for an extra unit of health benefit (QALYs or ulcer-free weeks in this case), B is the health benefits under treatment t, C is the cost under treatment t, t is an index for treatment type and θ is the set of unknown parameters of a decision model.

The estimated mean QALYs, ulcer-free weeks and total costs associated with placebo and pentoxifylline were combined with a feasible range of values for decision makers' willingness to pay (λ), to obtain distributions of net benefits at different levels of λ. The uncertainty surrounding the decision to adopt pentoxifylline as a cost-effective treatment at different levels of willingness to pay was represented in acceptability curves. Cost-effectiveness acceptability curves are a graphical representation of the probability of an intervention being cost effective (on the vertical axis) for a range of willingness-to-pay values λ (on the horizontal axis) associated with the health outcome of interest.[26]

Value-of-Information Analysis

The expected value of perfect information (EVPI) associated with the decision to recommend pentoxifylline as a cost-effective adjunct to compression in the treatment of venous leg ulcers was estimated at different levels of willingness to pay for an extra unit of health benefit (λ) using the following non-parametric method:[27]

- EVPI for an individual evaluating a set of alternative treatments (t = 1,2...,T) is as seen in equation 5:

$$EVPI = E_\theta \max_t NB(t,\theta) - \max_t E_\theta NB(t,\theta) \quad \text{(Eq. 5)}$$

Table III. Model stochastic parameters

Parameter	Mean estimate (95% CI)
Prior analysis	
Probability of healing (weekly basis)	
pentoxifylline	0.046 (0.034, 0.059)
placebo	0.025 (0.016, 0.036)
Posterior analysis	
Probability of healing (weekly basis)	
pentoxifylline	0.044 (0.036, 0.052)
placebo	0.027 (0.020, 0.036)
Common to both analyses	
Probability of re-occurring (weekly basis)	0.006 (0.005, 0.007)
Probability of dying (weekly basis)	0.0003 (0.0002, 0.0005)
Healing utility weight	0.75 (0.12, 0.99)
Open ulcer utility weight	0.64 (0.07, 0.99)
Death utility weight	0
Proportion on Setopress® [a] bandage	0.02 (0.01, 0.03)
Proportion on Surepress® bandage	0.03 (0.01, 0.04)
Proportion on Tensopress® bandage	0.13 (0.11, 0.16)
Proportion on Comprilan® bandage	0.01 (0.005, 0.02)
Proportion on Ultra Four® bandage	0.01 (0.005, 0.02)
Proportion on Rosidal K® bandage	0.01 (0.008, 0.01)
Proportion on four-layer bandage	0.58 (0.55, 0.62)
Proportion on hosiery class 2/3	0.08 (0.06, 0.10)
Proportion on other bandages	0.13 (0.11, 0.15)
Proportion on NA dressing	0.51 (0.47, 0.54)
Proportion on hydrocolloid dressing	0.14 (0.12, 0.17)
Proportion on foam dressing	0.06 (0.04, 0.08)
Proportion on alginate dressing	0.05 (0.04, 0.07)
Proportion on tulle grass dressing	0.02 (0.01, 0.03)
Proportion on other dressing	0.22 (0.19, 0.25)
Proportion of grade D nurses	0.04 (0.02, 0.05)
Proportion of grade E nurses	0.47 (0.43, 0.52)
Proportion of grade F nurses	0.04 (0.03, 0.06)
Proportion of grade G nurses	0.44 (0.40, 0.48)
Proportion of grade H nurses	0.01 (0.003, 0.02)
Proportion visited once a week	0.67 (0.63, 0.70)
Proportion visited twice a week	0.23 (0.21, 0.27)
Proportion visited thrice a week	0.10 (0.08, 0.12)
Proportion visited for 20 minutes	0.35 (0.31, 0.38)
Proportion visited for 30 minutes	0.51 (0.48, 0.55)
Proportion visited for 40 minutes	0.14 (0.11, 0.16)
Proportion visited at home	0.47 (0.43, 0.50)
Proportion visited at a clinic	0.53 (0.50, 0.57)

a The use of trade names is for product identification purposes only and does not imply endorsement.

CI = credibility interval; **NA** = knitted viscose primary dressing.

where E_θ = expectation over θ, $\max_t$ = maximum over t, t = fixed number of treatments (t = 1,2...,T) and θ = unknown number of parameters. Thus, $E_\theta \max_t NB(t,\theta)$ is the mean of the maximum net benefit for each iteration (assumes perfect information), $\max_t E_\theta NB(t,\theta)$ is the mean net benefit of the treatment with the maximum mean net benefit over all iterations,

- For the entire population of individuals with open venous leg ulcers, this translates into the population EVPI (PEVPI) as per equation 6:

$$PEVPI = EVPI \sum_{i=1}^{I} \frac{P_i}{(1+r)^i} \quad \text{(Eq. 6)}$$

where P_i = incidence of open venous leg ulcers at time i (i = 1,., I), r = discount rate, I = effective lifetime of the new technology.

PEVPI relates to all patients who could potentially benefit from additional research over the lifetime of the new technology.

EVPI was estimated in both the prior and posterior analyses. In both stages of the analysis, a 10-year life span for the research findings was assumed, and the annual expected value of perfect information was discounted at 6%. To date, no data are available regarding the incidence of venous leg ulcers; this parameter was estimated using data on the prevalence of venous leg ulcers in the UK general population aged ≥18 years (1.5 per 1000 people). It was also assumed that <20% of prevalent venous leg ulcers are open at any time, 40% of which will not heal with compression treatment after 6 months and may require adjunct therapy to compression to achieve healing.

Results

Mean estimates of all stochastic and deterministic model parameters are described in table III.

Prior Analysis

In the first stage of the analysis, pentoxifylline was identified as a dominant therapy (i.e. in comparison with placebo, the administration of this drug produced greater health benefits at a lower cost). People in the pentoxifylline group healed an average of 8.28 weeks quicker (95% CI 1.89, 14.56) than

people in the placebo group, had a 0.02 gain in QALYs (95% CI −0.12, 0.17) and an average reduction in cost of £153.4 (95% CI −53.11, 354.9) [figure 2].

In the base-case analysis, the distribution of the incremental costs and incremental QALYs in the cost-effectiveness plane (figure 3) suggests substantial uncertainty is associated with the dominance of pentoxifylline. The largest proportion of point estimates concentrate in the southwest and southeast quadrants. This indicates that, in a large proportion of the iterations, the mean differential in costs favours pentoxifylline; however, the mean differential in health benefits can go in either direction. In the southwest quadrant, cost savings are achieved at the expense of reducing the mean gain in QALYs, and in the southeast quadrant, pentoxifylline is both cheaper and produces greater health benefits (i.e. is a 'dominant alternative').

The decision uncertainty associated with the dominance of pentoxifylline is represented in figure 4. For a large range of considerably low willingness-to-pay values for an extra QALY gained, the acceptability curve for pentoxifylline suggests that this therapy is associated with a high probability of being cost effective in the treatment of chronic venous leg ulcers. For instance, a 93% probability of being cost effective was estimated for pentoxifylline at a willingness-to-pay value of £50 for an extra QALY.

The Bayesian value-of-information analysis, based on the data from the three RCTs[11-13] published prior to the Dale et al.[10] study, suggested that, given the considerably low levels of uncertainty associated with the decision to adopt pentoxifylline as a cost-effective treatment, the acquisition of further primary data may not be justified (figure 5). For willingness-to-pay values of £0, £100 and £500 per QALY gained, the EVPIs were £128 200, £127 100 and £126 700, respectively. The approximate total cost of another trial on compression for venous leg ulcers, started in 1998 and completed in 2001, was £380 000.[25,28] Cross referencing this value with our estimates of the maximum expected return value associated with primary research (i.e. 'value of perfect information'), it is only at willingness-to-pay values above £1500 for an extra QALY that the collection of further data to reduce the decision uncertainty associated with the adoption of pentoxifylline would have been justified. This analysis suggests that, unless decision makers were willing to pay £1500 or more per extra QALY for people with chronic venous leg ulcers in 1998, the actual cost of primary research may have exceeded the expected return value associated with the information provided by this study.[10]

Posterior Analysis

In the second analysis, the clinical effectiveness data from Dale et al.[10] were used to ascertain whether or not the prior belief that this trial was unlikely to be a 'cost effective' use of research resources was correct.

Data from Dale et al.,[10] added to the data from the three earlier RCTs,[11-13] were used to update the weekly probabilities of healing. The results regarding the cost effectiveness of pentoxifylline based on these posterior probabilities were slightly modified. In the posterior analysis, pentoxifylline was associated with a mean expected cost saving of £98.09 (95% CI −49.21, 245.00), a mean expected QALY gained of 0.01 (95% CI −0.09, 0.13) and a reduction in time to healing of 6.55 weeks (95% CI 1.93, 11.15). On average, pentoxifylline was still a dominant strategy compared with placebo. However, as the results in figure 2 indicate, the point estimates of the mean incremental costs and health benefits were smaller than those observed in the prior analysis.

The distribution of differential costs and differential QALYs in the cost-effectiveness plane followed a similar pattern to that observed in the prior analysis; however, the point estimates were more precise in the posterior analysis (figure 3). The distribution of points in the posterior cost-effectiveness plane shrank, reflecting the reduction in the posterior variance of the measure of health benefit as a result of additional data from Dale et al.[10] However, while the posterior variance is smaller, the upward shift of the posterior cost-effectiveness plane results in greater overall decision uncertainty.

Relative to the prior analysis, the probability of pentoxifylline being cost effective was slightly reduced for the same range of willingness-to-pay values. Estimated posterior probabilities of pentoxifylline being cost effective at willingness-to-pay values

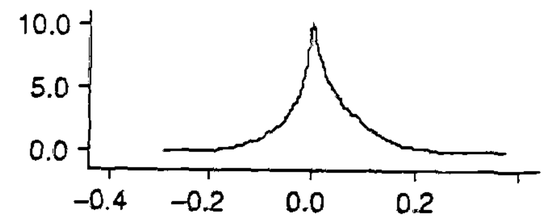
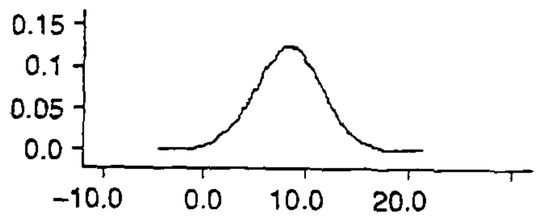
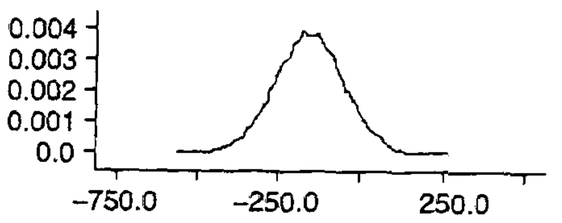
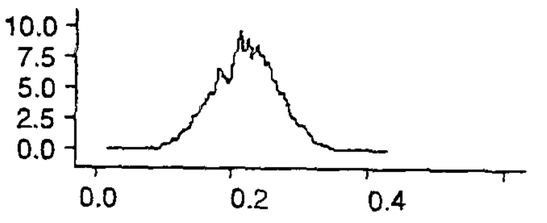
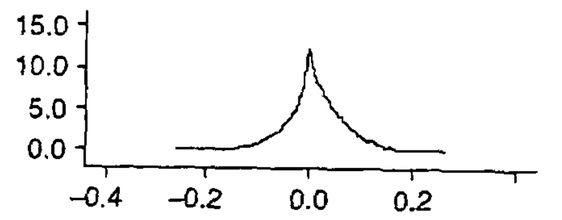
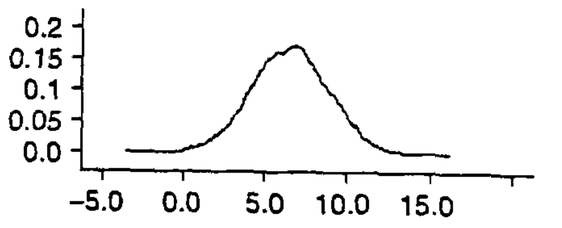
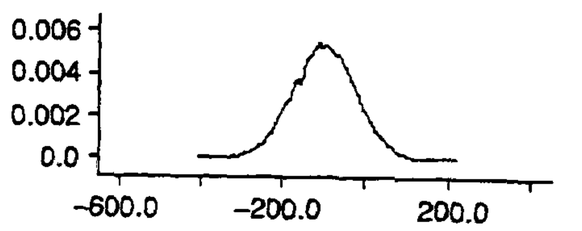
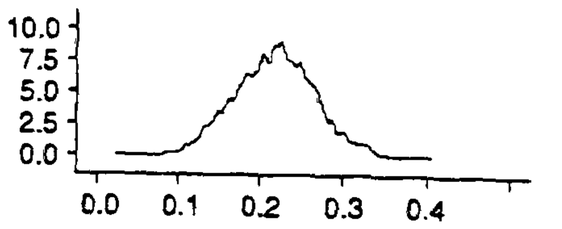
	Mean (95% credibility interval)	Posterior distribution
Prior analysis		
Incremental QALYs	0.02 (–0.12, 0.17)	
Incremental ulcer-free weeks	8.28 (1.89, 14.56)	
Incremental cost	£153.4 (–53.11, 354.9)	
Heterogeneity between studies	0.22 (0.13, 0.31)	
Posterior analysis		
Incremental QALYs	0.01 (–0.09, 0.13)	
Incremental ulcer-free weeks	6.55 (1.93, 11.15)	
Incremental cost	£98.09 (49.21, 245.0)	
Heterogeneity between studies	0.22 (0.12, 0.32)	

Fig. 2. Main results of the prior and posterior incremental cost-effectiveness analysis (2004 values, UK NHS perspective) of pentoxifylline (vs placebo) in addition to compression. For the posterior distributions, each y-axis represents frequency and each x-axis represents the relevant parameter (incremental QALYs, ulcer-free weeks or costs or heterogeneity).

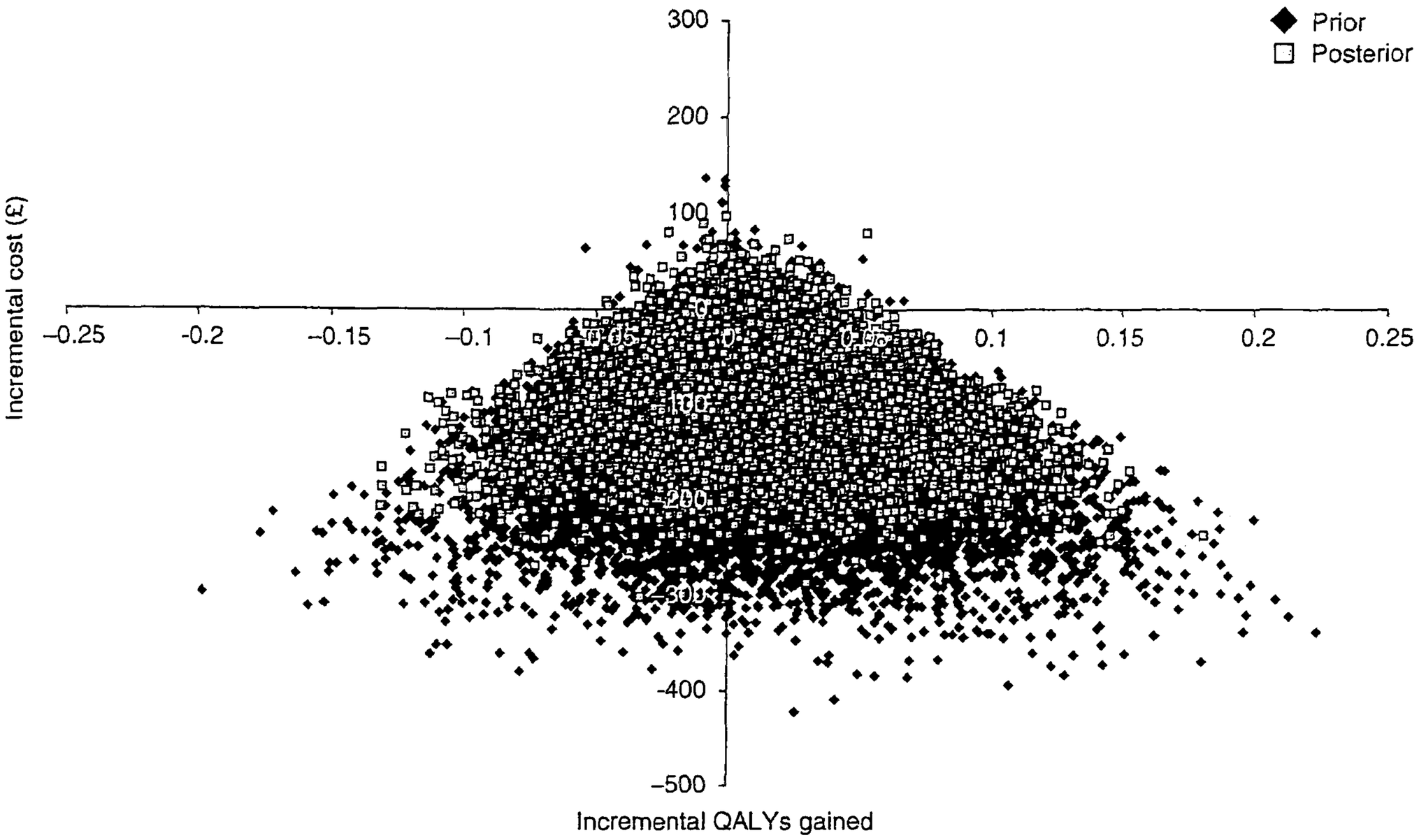


Fig. 3. Cost-effectiveness plane showing the distribution of incremental costs (2004 values, UK NHS perspective) and QALYs for pentoxifylline vs placebo in addition to compression.

of £0, £500 and £1000 per QALY gained were 0.90, 0.91 and 0.89, respectively (figure 4). This increase in the level of decision uncertainty surrounding the cost effectiveness of pentoxifylline may be due to the fact that, by incorporating more information into the analysis, we not only reduced the size of the mean incremental health benefit associated with pentoxifylline but also changed the distribution of the mean incremental costs, thus significantly modifying the distribution of the posterior mean expected monetary benefit.

The results from the posterior value of perfect information analysis depicted in figure 5 followed the same direction as those from the prior analysis. Low posterior EVPI values were associated with willingness-to-pay values in the range of £0 to £2500 per QALY gained. A comparison of the posterior EVPI values in figure 5, with the same cut-off value of £380 000 used in the prior analysis, suggested once again that unless a decision maker's willingness-to-pay values for an extra QALY gained are greater than £1500, the acquisition of further information is unlikely to be an efficient use of resources.

Cost-Effectiveness Analysis

Variations in the definition of health benefit (i.e. ulcer-free weeks rather than QALYs), yielded similar results to those from the base-case cost-utility analysis. The use of natural units merely emphasised the 'dominance' of pentoxifylline as an adjunct therapy and reduced the decision uncertainty associated with its adoption, that is, the EVPI was reduced for the same range of willingness-to-pay values for an extra unit of health benefit used in the base-case analysis (these data are not presented but can be requested from the corresponding author).

Sensitivity Analysis

Evidence of the 'dominance' of pentoxifylline as an adjunct therapy to compression (vs placebo) in the treatment of chronic venous leg ulcers was provided in all the sensitivity analyses. The model was robust with regards to the length of follow-up for effectiveness data, and the prior distributions assigned to pooled estimates of the natural logarithm of the odds ratio of healing or to heterogeneity between studies.

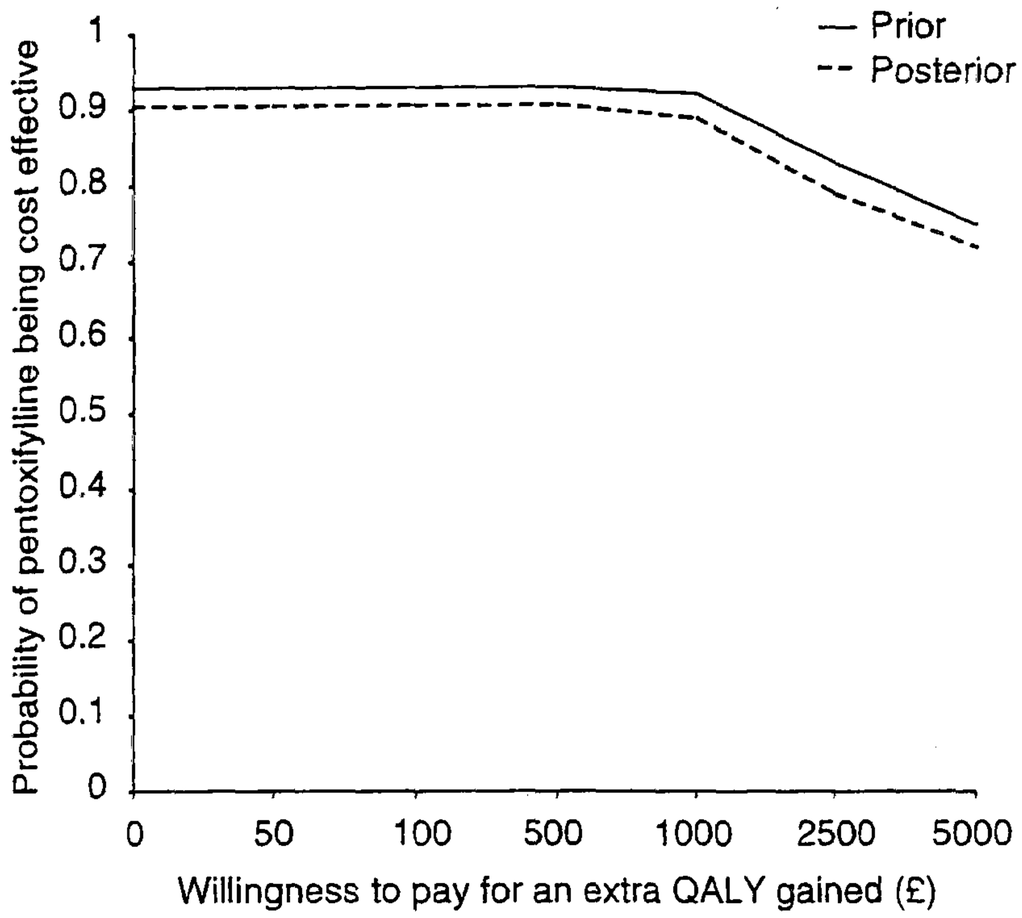


Fig. 4. Cost-effectiveness acceptability curve (prior analysis) showing the decision uncertainty around the dominance of pentoxifylline vs placebo (2004 values) from the UK NHS perspective.

The EVPI was significantly influenced by variations in the definition of the parameter describing the heterogeneity between studies and the length of follow-up. The greater degree of heterogeneity between studies that was assumed, the higher the values associated with the collection of further primary information. As previously mentioned, a small number of studies is not sufficient to precisely estimate the heterogeneity between studies, thus considerable judgement was required to define a 'realistically vague prior distribution'[21] (these data are not presented but can be requested from the corresponding author).

Discussion

In ascertaining the clinical and cost effectiveness of healthcare technologies, primary research is one of the most valuable and reliable sources of information for decision making. However, the scarcity of resources available for healthcare research and development (R&D), requires decision makers to confront their need for further information with the possibility of substantiating their decision on existing information, and thus freeing resources that can be allocated to other areas of priority.

A Bayesian value-of-information analysis explicitly quantifies the levels of uncertainty associated with a decision at a specific point in time and given a fixed amount of existing information. Whilst the collection of further information can reduce the degree of uncertainty inherently associated with a decision, it has been argued that R&D resources may be more efficiently allocated if primary research is substantiated only when the expected reduction in the levels of uncertainty attributable to the newly acquired information are likely to counterbalance the actual cost of research.[10]

The promotion of pentoxifylline as an adjunct pharmaceutical therapy to compression in the treatment of venous leg ulcers has been largely prevented because of the equivocal findings reported in the literature regarding this drug's ability to expedite the healing process of chronic venous ulcers. A series of RCTs of different sizes and designs have been conducted to ascertain the benefits associated with pentoxifylline.[9] However, none of these studies were considered to provide conclusive evidence regarding the clinical effectiveness of this drug. Furthermore, to the best of our knowledge, the cost effectiveness of pentoxifylline had not previously been formally explored.

Our findings suggest that the current degree of uncertainty associated with the use of pentoxifylline as an adjunct to compression in the treatment of chronic venous leg ulcers does not justify the conduct of further research. Furthermore, the Bayesian value-of-information analysis conducted using data

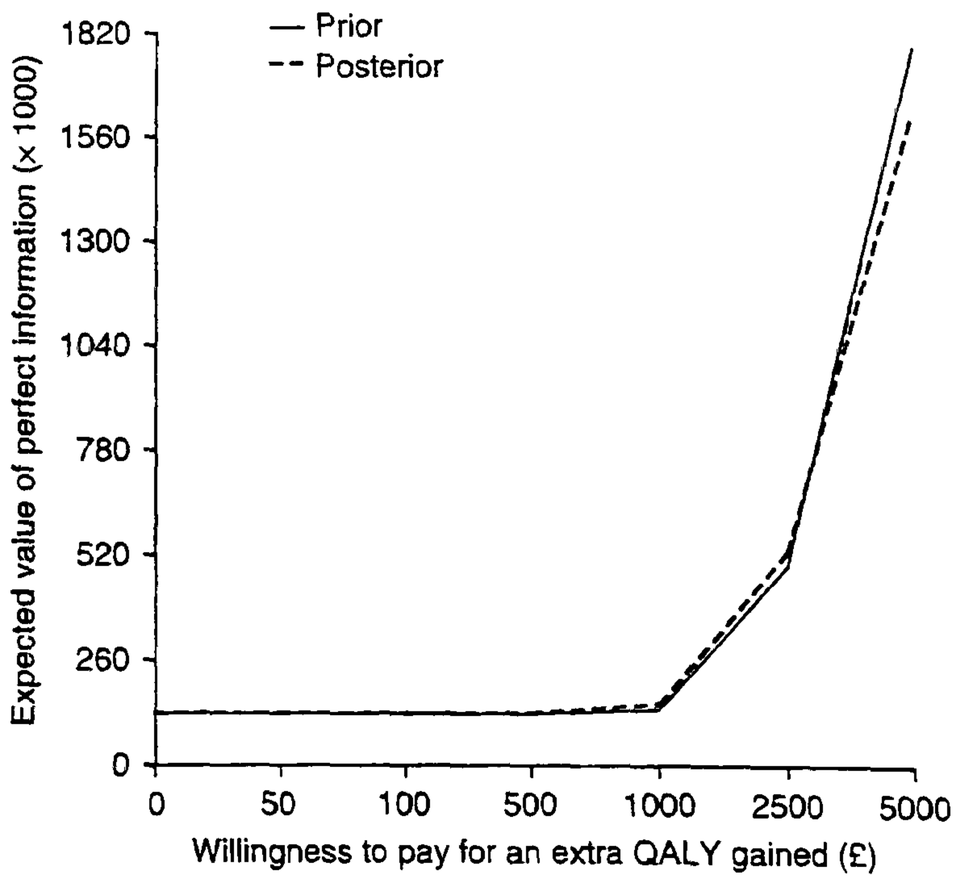


Fig. 5. Expected value of perfect information (2004 values) on the cost effectiveness of pentoxifylline (vs placebo) from the UK NHS perspective, according to a range of decision maker's willingness to pay for an extra QALY gained from £0 to £5000.

from the three studies[11-13] published before the UK-based largest study in this area (prior analysis),[10] suggested that pentoxifylline was already associated with a large probability of being not only clinically but also cost effective, if not cost saving.

Whilst the Dale et al.[10] study on pentoxifylline contributed to a more precise estimate of its likely associated benefit, its inclusion in the estimate of the pooled odds ratio of healing at 24 weeks did not reverse the conclusion from the prior analysis that pentoxifylline was likely to be a cost-effective intervention. Had the results from the study by Falanga et al.[13] and the framework proposed by Claxton et al.[4] been available prior to the commissioning of the study by Dale et al.,[10] a more efficient use of R&D resources could have been made.

Interestingly, two further trials[29,30] reporting a statistically significant beneficial effect of pentoxifylline have been published since the trial by Dale et al.[10] These two Italian/British studies were funded with grants awarded from healthcare institutions without any support from the pharmaceutical industry. It is intriguing that, despite the mounting evidence on the clinical and cost effectiveness of pentoxifylline for the treatment of venous leg ulcers, there is a lack of initiative from its manufacturers to have this drug's licence amended. In addition to the direct health benefit to people with chronic venous leg ulcers, the official recognition of pentoxifylline's indication as an adjunct to compression therapy in the treatment of venous leg ulcers may prevent further inefficient use of public resources that are being used worldwide to conduct primary research on the effectiveness of this drug.

This analysis provides evidence of the valuable potential contribution of Bayesian value-of-information analysis to assist the decision-making process in healthcare.[4] This methodology not only facilitates the integration of existing evidence in a coherent structure, but also allows estimation of the level of uncertainty associated with decisions regarding the adoption of new healthcare technologies. In this analysis, the framework was further strengthened by making use of a comprehensive analytic-decision model to evaluate the cost effectiveness of pentoxifylline compared with placebo. This enabled us to make full allowance for potential inter-relationships between model parameters, incorporate greater parameter uncertainty by allowing us to estimate the pooled population effect and any potential baseline heterogeneity between studies from the data and make a more adequate stochastic characterisation of existing data without unnecessarily recurring to assumptions of normality.[5]

The current analysis only considered evidence from RCTs up to the publication of the findings from Dale et al.[10] While this can be considered as a limitation, we decided to do so to exemplify the potential contribution of the proposed framework at that specific point in time and given the set of evidence available then. Only a few RCTs have investigated the effect of pentoxifylline as an adjunct therapy to compression for treating chronic venous leg ulcers. Therefore, considerable judgement was required to characterise the prior distribution of the parameter describing the potential heterogeneity between studies. While evidence regarding the clinical effectiveness of pentoxifylline from a wider range of clinical study designs could have been included, such analysis would have required considerably more human and financial resources than we had available.

Conclusion

The available information on pentoxifylline suggests that this drug is associated with a high probability of being cost effective in the treatment of chronic venous leg ulcers. Future research in this area can now concentrate on ensuring an optimal use of this therapeutic agent, for example by evaluating alternative methods of administration, establishing an optimal dosage, identifying the patient group most likely to benefit and estimating the point in time when an adjunct therapy should be considered.

Acknowledgements

Cynthia P. Iglesias conceived the idea, conducted the analyses and was responsible for the preparation of this manuscript. Karl Claxton contributed to the idea, provided expert advice on decision-analytic modelling and value-of-perfect-information analysis and commented on previous drafts of this manuscript.

This work was funded with a pump priming research grant from the University of York.

We would like to thank Andrea Manca, David Torgerson and Nicky Cullum for commenting on previous versions of this manuscript, and Christian Asseburg for his advice on

Bayesian meta-analysis, as well as the two anonymous referees for their helpful comments. Finally, we would like to thank Clare Morrell and the Royal College of Nursing for allowing us to use data from the Royal College of Nursing's Sentinel Clinical Audit Dataset.

The authors have no conflict of interest directly relevant to the content of this article.

References

1. Claxton K. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. J Health Econ 1999; 18: 341-64
2. Stinnet AA, Mullay J. Net health benefits: a new framework for the analysis of uncertainty in cost-effectiveness analysis. Med Decis Making 1998; 18: S68-80
3. Fleurence RL, Torgerson DJ. Setting priorities for research. Health Policy 2004; 69: 1-10
4. Claxton K, Sculpher M, Drummund M. A rational framework for decision making by the National Institute for Clinical Excellence (NICE). Lancet 2002; 360: 711-5
5. Cooper NJ, Sutton AJ, Abrahms KR, et al. Comprehensive decision analytical modelling in economic evaluation: a Bayesian approach. Health Econ 2000; 13: 203-26
6. Callam MJ, Ruckley CV, Harper DR, et al. Chronic ulceration of the leg: extent of the problem and provision of care. BMJ 1985; 290: 1855-6
7. Bosanquet N. Costs of venous ulcers: from maintenance therapy to investment programmes. Phlebology 1992; 1: 44-6
8. Brenner MA. Non healing venous stasis ulcers: pentoxifylline as adjunctive therapy. J Am Podiatr Med Assoc 1987; 77: 586-8
9. Jull AB, Waters J, Arrow B. Pentoxifylline for treatment of venous leg ulcers: systematic review. Lancet 2002; 359: 1551-4
10. Dale JJ, Ruckley CV, Harper DR, et al. Randomised, double blind placebo controlled trial of pentoxifylline in the treatment of venous leg ulcers. BMJ 1999; 319: 875-8
11. Schurmann W, Eberhardt R. The efficacy of pentoxifylline added to topical and compression therapy in patients with varicose and postthrombotic leg ulcers. Therapiewoche 1986; 36: 2343-5
12. Colgan MP, Dormandy J, Jones P, et al. Oxpentifylline treatment of venous ulcers of the leg. BMJ 1990; 300: 972-5
13. Falanga V, Fujitani RM, Diaz C, et al. Systemic treatment of venous leg ulcers with high doses of pentoxifylline: efficacy in a randomized, placebo-controlled trial. Wound Repair Regen 1999; 7: 208-13
14. Government Actuary's Department. Interim life tables 2003-based. 2004 [online]. Available from URL: http://www.gad.gov.uk/life_tables [Accessed 2005 May]
15. Franks PJ, Oldroyd MI, Dickson D, et al. Risk factors for leg ulcer recurrence: a randomized trial of two types of compression stocking. Age Ageing 1995; 24: 490-4
16. Morrell C, Liao XH, Cheater F, et al. The management of venous leg ulcers: a project to improve care. Nurs Stand 2001 Apr 11-17; 15 (30): 68-73
17. British National Formulary (September). British Medical Association/Royal Pharmaceutical Society of Great Britain. Vol. 40. London: BNF, 2004
18. Booth C. Review body for nursing staff, midwifes, health visitors and professions allied to medicine. London: HMSO, 2003 [online]. Available from URL: http://www.ome.uk.com/review.cfm?.body=6.20(Cm5716) [Accessed 2005 May]
19. HM Treasury. Appraisal and evaluation in central government: treasury guidance [online]. Available from URL: http://www.hm-treasury.gov.uk [Accessed 2005 May]
20. Spiegelhalter DJ, Thomas A, Best NG, et al. WinBUGS version 1.4 user manual. Cambridge: MRC Biostatistics Unit, 2002 [online]. Available from URL: www.mrc-bsu.cam.ac.uk/bugs [Accessed 2002 Apr]
21. Lambert PC, Sutton AJ, Burton PR, et al. How vague is vague? A simulation study of the impact of the use of vague prior distributions in MCMC using WinBUGS. Stat Med 2005 Aug 15; 24 (15): 2401-28
22. Spiegelhalter DJ, Abrams KR, Myles JP. Bayesian approaches to clinical trials and health-care evaluation. Chichester: John Wiley & Sons Ltd, 2004
23. Spiegelhalter DJ, Best NG. Bayesian methods for evidence synthesis and complex cost-effectiveness models: an example in hip prostheses. Stat Med 2003; 22: 3687-709
24. Gelman A, Carlin JB, Stern HS, et al. Bayesian data analysis. 2nd ed. Boca Raton (FL): Chapman & Hall/CRC, 2004
25. Iglesias CP, Nelson EA, Cullum NA, et al. VenUS: a randomised controlled trial of two bandages for treating venous leg ulcers. Health Technol Assess 2004; 8 (24): 41-8
26. Van Hout VA, Ai MJ, Gordon GS, et al. Cost, effects, and C/E ratios alongside a clinical trial. Health Econ 1994; 3: 309-19
27. Ades AE, Lu G, Claxton K. Expected value of sample information calculations in medical decision modeling. Med Decis Making 2004; 24: 207-27
28. UK National Health Service Health Technology Assessment programme [online]. Available from URL: http://www.ncchta.org/ProjectData/3_project_record_published.asp?PjtId=1076&Searchtext=varicose+veins [Accessed 2005 May]
29. Belcaro G, Cesarone MR, Nicolaides AN, et al. Treatment of venous ulcers with pentoxifylline: a 6-month randomized double-blind, placebo controlled trial. Angiology 2002; 53: S45-7
30. De Sanctis MT, Belcaro G, Cesarone MR, et al. Treatment of venous ulcers with pentoxifylline: a 12-month double-blind, placebo controlled trial: microcirculation and healing. Angiology 2002; 53: S49-51

Correspondence and offprints: *Cynthia P. Iglesias*, Department of Health Sciences, University of York, Seebohm Rowntree Building, Area 2, York, YO10 5DD, UK.
E-mail: cpiu1@york.ac.uk

Appendix B

Methodological issues associated with clinical evidence

Some of the most frequently reported methodological limitations associated with clinical studies in VLUs are:[7-9]

i) *Inadequate comparisons*: for example, a recently published study reported a comparison between *4-layer* bandage and "other available treatments", i.e. any type of compression and wound dressing indicated for VLUs.[17] Given the large number of factors that could have systematically influenced the results observed in the control group, any effect observed in this group cannot be reliably attributed to any specific treatment alternative. Inadequate comparisons like this not only introduce a considerable potential for confounding, but just as importantly, represent an inefficient use of scarce R&D resources.

ii) *Small samples reported without sample size calculations*: RCTs with between 12 and 25 participants are common in VLUs.[7-9] For example, the only RCT in the international literature reporting on the effectiveness of larval therapy as a treatment for VLUs had only 12 participants.[61] Small samples reduce the statistical power of the analysis to draw any reliable inferences from the data by increasing the probability of a type II error, i.e. not rejecting the null hypothesis when the null hypothesis is not true.

iii) *Inadequate duration of follow-up*: RCTs reporting on the effectiveness of pentoxifylline as an adjunct to compression for the treatment of VLUs followed up

participants for between 8 and 24 weeks.[40] While between 30% - 60% of VLUs would heal within this period, there is a large proportion of ulcers that would fail to do so.[128] In studies with short duration of follow-up there will be an artificially small proportion of participants achieving the endpoint, in this case healing, hence the statistical power of the study could be compromised. Furthermore, the recurrent nature of VLUs highlights the relevance of monitoring individuals over longer periods of time. Resolving an isolated episode of ulceration is only one of the objectives of VLU management, the principal aim of which should be to maximise ulcer free time. This is consistent with a decision analytic approach to evaluation, which suggests that the relevant time horizon for evaluation should be that over which the health condition under study is likely to be active, and over which cost and health outcomes are expected to differ.[95] In the case of chronic conditions this will be individual's life time.

iv) *Imbalances between trial arms in participant's baseline characteristics*: area and duration of current VLUs have been identified as predictors for VLU healing.[129;130] A RCT of long-stretch versus *short-stretch* compression bandages reported baseline imbalances both in the median ulcer size and median ulcer duration.[131] In spite of this, the statistical analysis of this study did not include an adjustment for baseline covariates, thus potentially reporting a biased estimate of the treatment effect.[132] This problem of baseline imbalance is a common one in wound care studies generally.[133]

v) *The use of intermediate* (reductions in area of ulceration) *and sub-optimal outcomes* (simple proportions of ulcers healed rather than time to healing) over an arbitrarily fixed period of time is frequently reported in wound care RCTs.[58;134] Whilst changes in ulcer shape and size are common, ulcers may still remain unhealed

for considerable periods of time, hence effectiveness claims based on this type of outcome can be highly misleading. Similarly, simple proportions of ulcers healed at a specific point in time provide partial and arbitrary information about the actual effect associated with any therapy.[16] The most informative outcome from a clinical and economic perspective will be "time to ulcer free person".

Methodological issues associated with economic evidence

Frequently encountered shortcomings associated with economic evaluation studies on VLUs:

i) *Use of uninformative intermediate outcomes*: in cases where the measurement of final outcomes is precluded, the use of intermediate outcomes may be justified, on the basis of a strong association between intermediate and final outcomes. However, rates of change in ulcer size are not associated in a known way with time to ulcer healing, therefore intermediate outcomes based on changes in area of ulceration may be misleading.[58;134]

ii) *Estimates of cost only based on unit cost of treatments*: in an earlier VLU economic evaluation analysis the only component of cost considered was the "unit cost" of the competing alternative treatments.[135] This approach for the estimation of total costs is not only incomplete but also highly misleading since overall treatment costs are made out of a number of items that may contribute differently to estimate overall treatment costs. Nursing time, i.e., time nurses dedicate to looking after patients, has been identified as the greatest component of VLU treatment.[52] Thus, unless considerable reductions in nursing time can be associated with a treatment, a differential in purchasing cost does not directly imply "cost-effectiveness".

iii) *Use of medians rather than means*: estimates of the median rather than mean total treatment cost have been used in some analyses.[124;135] Total treatment cost data usually follow positively skewed distributions, i.e., data distributes asymmetrically. While the median may be more appropriate to describe a "typical" observation,[84] budgetary decisions and decisions about which intervention is worthwhile need to be based on what the cost and outcomes are expected to be, i.e. the average of the population instead of on only 50% of it.[136;137] Consequently, economic evaluation analyses need to be based on estimates of the population mean otherwise the budget impact would be under estimated and estimates of cost-effectiveness will be biased since in positively skewed distributions the median is smaller than the mean.

Appendix C

Databases: CINAHL, MEDLINE, EMBASE

Search Strategy:

---

1 skin ulcer/ or leg ulcer/ or pressure ulcer/ (1325)

2 exp Foot Ulcer/ (647)

3 exp Diabetic Neuropathies/ (773)

4 Diabetic Angiopathies/ (31)

5 Diabetes Mellitus/co [Complications] (467)

6 diabetes mellitus, insulin-dependent/co (107)

7 diabetes mellitus, non-insulin-dependent/co (183)

8 Pilonidal Cyst/ (11)

9 Surgical Wound Infection/ (301)

10 ((plantar or diabetic or heel or venous or stasis or arterial) adj4 ulcer$).ti,ab. (433)

11 ((decubitus or foot or diabetic or ischaemic or pressure) adj4 ulcer$).ti,ab. (732)

12 ((pressure or bed) adj4 sore$).ti,ab. (279)

13 (pilonidal adj4 (cyst or sinus)).ti,ab. (6)

14 (bedsore or bedsores or (bed adj sore$)).ti,ab. (22)

15 ((diabetic adj foot) or (cavity adj4 wound$)).ti,ab. (304)

16 ((varicose or leg or skin) adj4 ulcer$).ti,ab. (396)

17 ((decubitus or chronic or sinus) adj4 wound$).ti,ab. (183)

18 ((burn or gunshot or bite) adj4 wound$).ti,ab. (159)

19 exp BURNS/ (957)

20 (diabetic adj (neuropath$ or angiopath$)).ti,ab. (88)

21 (diabetes adj4 (wound$ or ulcer$)).ti,ab. (88)

22 (surgical adj wound adj infection).ti,ab. (10)

23 ((dehiscen$ or sepsis or exudat$ or necrot$ or slough$) adj4 (wound$ or ulcer$)).ti,ab. (96)

24 (((non adj heal$) or nonheal$ or problem or difficult$ or complic$) adj4 (wound$ or cavity or incision$)).ti,ab. (174)

25 ((granulating or postoperative) adj wound$).ti,ab. (14)

26 exp dental health services/ or exp dentistry/ or exp peptic ulcer/ or exp corneal diseases/ (2838)

27 case study.pt. (13526)

28 or/26-27 (16003)

29 exp clinical trials/ or random assignment/ or placebos/ or meta analysis/ or exp prospective studies/ (18475)

30 systematic review/ or comparative studies/ or clinical trial.pt. or review.pt. or systematic review.pt. (22625)

31 (clinical adj trial$).ti,ab. (1859)

32 ((singl$ or doubl$ or trebl$ or tripl$) adj blind$).ti,ab. (1040)

33 (control or controls or controlled or controlling or metaanalys$).ti,ab. (16772)

34 (meta adj analys$).ti,ab. (688)

35 (random$ or prospective$ or (comparison adj group$) or (standard adj treatment$)).ti,ab. (11937)

36 (placebo$ or (systematic adj review$)).ti,ab. (2809)

37 or/29-36 (46602)

38 economics/ or exp "costs and cost analysis"/ or economic aspects of illness/ (4489)

39 economics, pharmaceutical/ or economic value of life/ (108)

40 exp "fees and charges"/ or budgets/ (1358)

41 (cost or costs or economic$ or pharmacoeconomic$ or price$ or pricing$).ti,ab. (8446)

42 or/38-41 (11768)

43 (trauma adj4 (wound$ or ulcer$)).ti,ab. (33)

44 (pilonidal adj abscess$).ti,ab. (0)

45 or/1-25,43-44 (4624)

46 45 not 28 (4027)

47 46 and 37 (1201)

48 46 and 42 (269)

49 from 47 keep 1-1201 (1201)

50 from 49 keep 1-1201 (1201)

51 from 49 keep 1-1201 (1201)

52 from 49 keep 1-1201 (1201)

53 from 49 keep 1-1201 (1201)

54 from 49 keep 1-1201 (1201)

55 from 49 keep 1-1201 (1201)

56 from 48 keep 1-269 (269)

57 from 48 keep 1-269 (269)

58 from 48 keep 1-269 (269)

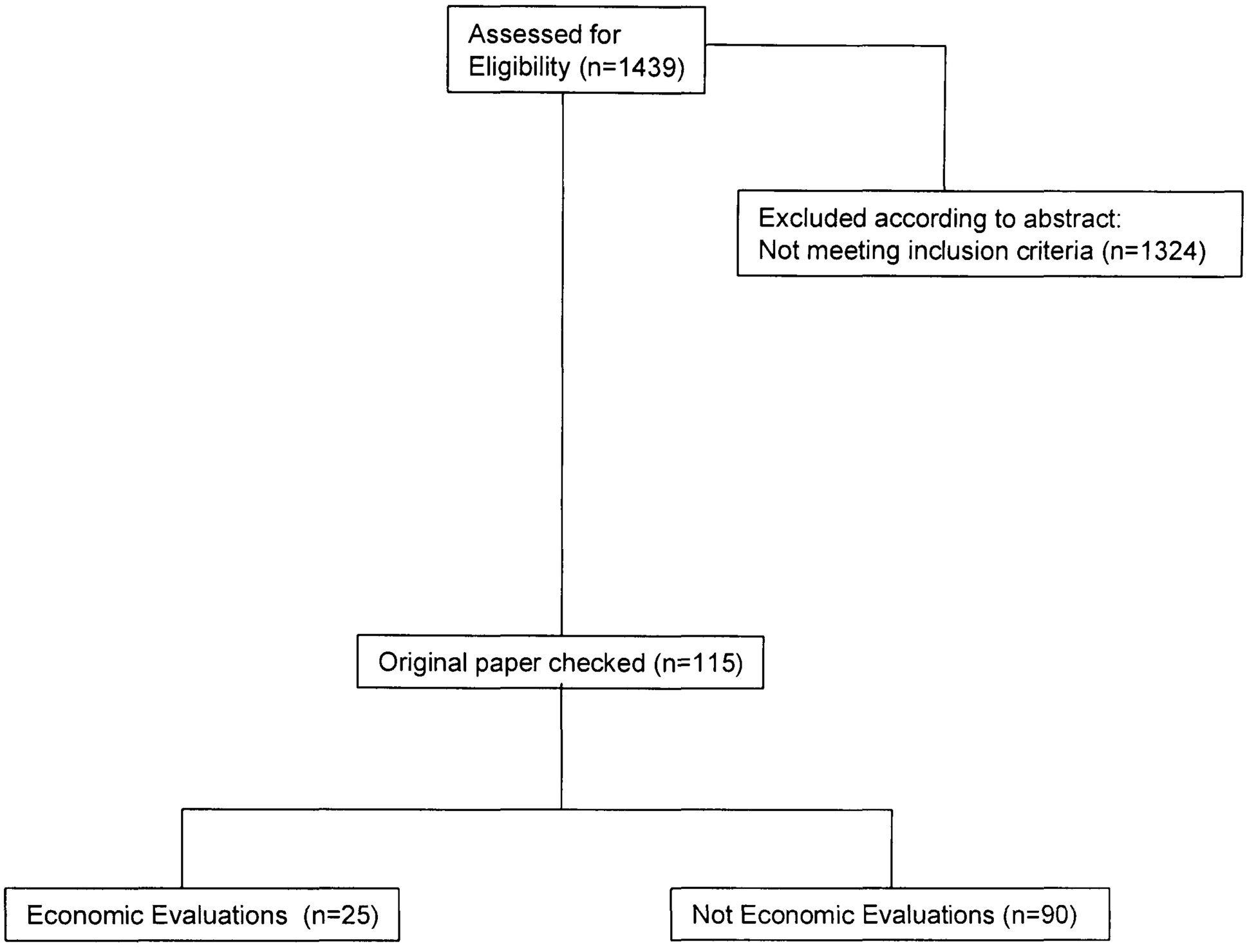


Figure 2. Quorum chart of economic evidence.[55]

Bibliography

1. Hjelmgren J, Berggren F, Andersson F. Health economic guidelines - similarities, differences and some implications. Value in Health 2001;4:225-50.

2. Commonwealth Department of Health Housing and Community Services. Guidelines for the Pharmaceutical Industry on Preparation of Submissions to the Pharmaceutical Benefits Advisory Committee. Canberra: AGPS; 1992.

3. Ministry of Health. Guidelines for Economic Analysis of Pharmaceutical Products. Toronto: Ontario Ministry of Health; 1994.

4. National Institute for Health and Clinical Excellence (NICE). Guide to the methods of technology appraisal. London: National Institute for Health and Clinical Excellence (NICE); 2004.

5. Medicines and Healthcare products Regulatory Agency. Guidance notes for manufactures on statistical considerations for clinical investigations of medical devices. MHRA; 2002. Report No.: 17.

6. Hennekens CH, Burke J. Design Strategies in Epidemiologic Research. In: Mayrent S.L., editor. Epidemiology in Medicine. 1st ed. 1987. p. 16-29.

7. Bradley M, Cullum N, Nelson EA, Petticrew M, Sheldon T, Torgerson D. Systematic reviews of wound care management: (2). Dressings and topical agents used in the healing of chronic wounds. Health Technol Assess 1999;3(17 Pt 2):1-35.

8. Cullum N, Nelson EA, Flemming K, Sheldon T. Systematic reviews of wound care management: (5) beds; (6) compression; (7) laser therapy, therapeutic ultrasound, electrotherapy and electromagnetic therapy. Health Technol Assess 2001;5(9):1-221.

9. O'Meara S, Cullum N, Majid M, Sheldon T. Systematic reviews of wound care management: (3) antimicrobial agents for chronic wounds; (4) diabetic foot ulceration. Health Technol Assess 2000;4(21):1-237.

10. Drummond MF, O'Brien BJ, Stoddart G, Torrance G. Methods for the Economic Evaluation of Health Care Programmes. 2nd ed. Oxford: Oxford University Press; 1997.

11. Drummond MF, Sculpher MJ, Torrance G, O'Brien BJ, Stoddart G. Methods for the Economic Evaluation of Health Care Programmes. 3nd ed. Oxford: Oxford University Press; 2005.

12. Philips Z, Ginnelly L, Sculpher MJ, Claxton K, Golder S, Riemsma R, et al. Review of guidelines for good practice in decision-analytic modelling in health technology assessment. Health Technology Assesement 2004;8(36):1-188.

13. National Institute for Health and Clinical Excellence (NICE). Technical Guidance for Manufacturers and Sponsors on making a Submission to a

Technology Appraisal, National Institute for health and Clinical Excellence. London: National Institute for Health and Clinical Excellence (NICE), 2001.

14. Sheldon TA. Health Technology Assessment. In: Armitage P, Colton T, editors. Encyclopedia of Biostatistics. 2nd ed. Chichester: John Wiley & Sons, Ltd; 2005. p. 2325-30.

15. Sculpher M, Drummond M, Buxton M. The iterative use of economic evaluation as part of the process of health technology assessment. J Health Serv Res Policy 1997 Jan;2(1):26-30.

16. Fletcher A, Cullum N, Sheldon TA. A systematic review of compression treatment for venous leg ulcers. BMJ 1997 Sep 6;315(7108):576-80.

17. O'Brien JF, Grace PA, Perry IJ, Hannigan A, Moloney MC, Burke PE. Randomized clinical trial and economic analysis of four-layer compression bandaging for venous ulcers. British Journal of Surgery 2003;90:794-8.

18. Ukat A, Konig M, Vanscheidt W, Munter K-C. Short-stretch versus multilayer compression for venous leg ulcers: a comparison of healing rates. Journal of Wound Care 2003;12(4):139-43.

19. Partsch H, Damstra RJ, Tazelaar DJ, Schuller-Petrovic S, Velders AJ, de Rooij MJ, et al. Multicentre, randomised controlled trial of four-layer bandaging versus short-stretch bandaging in the treatment of venous leg ulcers. Vasa 2001 May;30(2):108-13.

20. Meyer FJ, Burnand KG, Lagattolla NRF, Eastham D. Randomized clinical trial comparing the efficacy of two bandaging regimens in the treatment of venous leg ulcers. British Journal of Surgery 2002;89:40-4.

21. Meyer FJ, McGuinness CL, Lagattolla NR, Eastham D, Burnand KG. Randomized clinical trial of three-layer paste and four-layer bandages for venous leg ulcers. British Journal of Surgery 2003;90(8):934-40.

22. Moffatt CJ, McCullagh L, O'Connor T, Doherty DC, Hourican C, Stevens J, et al. Randomised trial of four-layer and two-layer bandage systems in the management of chronic leg ulceration. Wound Repair and Regeneration 2003;11(3):166-71.

23. Walters SJ, Morrell CJ, Dixon S. Measuring health-related quality of life in patients with venous leg ulcers. Qual Life Res 1999 Jun;8(4):327-36.

24. Wissing U, Ek AC, Unosson M. A follow-up study of ulcer healing, nutrition, and life-situation in elderly patients with leg ulcers. J Nutr Health Aging 2001;5(1):37-42.

25. Franks PJ, Moffatt CJ. Health related quality of life in patients with venous ulceration: use of the Nottingham health profile. Qual Life Res 2001;10(8):693-700.

26. Dale JJ, Ruckley CV, Harper DR, Gibson B, Nelson EA, Prescott RJ. Randomised, double blind placebo controlled trial of pentoxifylline in the treatment of venous leg ulcers. British Medical Journal 1999;319:875-8.

27. Callam MJ, Harper DR, Dale JJ, Ruckley CV. Chronic ulcer of the leg: clinical history. Br Med J (Clin Res Ed) 1987 May 30;294(6584):1389-91.

28. Graham ID, Harrison MB, Nelson EA, Lorimer K, Fisher A. Prevalence of Lower-limb Ulceration: A Systematic Review of Prevalence Studies. Advances in Skin and Wound Care 2003;16(6):305-16.

29. Moffat CJ, Franks PJ, Doherty DC, Martin R, Blewett R, Ross F. Prevalence of Leg Ulceration in a London population. Q J Med 2004;97:431-7.

30. Callam MJ, Ruckley CV, Harper DR, Dale JJ. Chronic ulceration of the leg: extent of the problem and provision of care. Br Med J (Clin Res Ed) 1985 Jun 22;290(6485):1855-6.

31. Dale JJ, Callam MJ, Ruckley CV, Harper DR, Berrey PN. Chronic ulcers of the leg: a study of prevalence in a Scottish community. Health Bull (Edinb) 1983 Nov;41(6):311-4.

32. O'Brien JF, Grace PA, Perry IJ, Burke PE. Prevalence and aetiology of leg ulcers in Ireland. Irish Journal of Medical Science 2000;169(2):110-2.

33. Cornwall JV, Dore CJ, Lewis JD. Leg ulcers: epidemiology and aetiology. Br J of Surgery 1986;73:693-6.

34. Walker N, Rodgers A, Birchall N, Norton R, MacMahon S. Leg ulceration as long-term complication of deep vein thrombosis. J Vasc Surg 2003;38(6):1331-5.

35. Valencia IC, Falabella A, Kirsner RS, Eaglstein WH. Chronic venous insufficiency and venous leg ulceration. J Am Acad Dermatol 2001 Mar;44(3):401-21.

36. Browse NL, Burnand KG. The cause of venous ulceration. Lancet 1982 Jul 31;2(8292):243-5.

37. Cullum NA, Roe B. Leg ulcers: nursing management. 1st ed. Middlesex: Scutari Press; 1995.

38. British Medical Journal. Clinical Evidence online www.clinicalevidence.org [Accessed May 2005]. 2005.

39. Cullum NA, Luker K, McInnes E, Nelson EA, Noakes H. The management of patients with venous leg ulcers: clinical practice guidelines. London: RCN Institute; 1998.

40. Jull AB, Waters J, Arrow B. Pentoxifylline for treatment of venous leg ulcers: systematic review. Lancet 2002;359:1551-4.

41. British Medical Association and Royal Pharmaceutical Society of Great Britain. British National Formulary. BMA and RPS of GB; 2005.

42. Franks PJ, Oldroyd MI, Dickson D, Sharp EJ, Moffatt CJ. Risk factors for leg ulcer recurrence: A randomized trial of two types of compression stocking. Age and Ageing 1995;24(6):490-4.

43. Harper DR, Ruckley CV, Gibson B, Brown D, Prescott RJ. Randomised trial of two grades of compression stockings in the prevention of venous ulcer recurrence - 5 year outcomes. Venous Forum 1999;14:91.

44. Spitzer WO. State of science 1986: quality of life and functional status as target variables for research. J Chronic Dis 1987;40(6):465-71.

45. McSweeny AJ, Creer TL. Health-related quality-of-life assessment in medical care. Dis Mon 1995 Jan;41(1):1-71.

46. Gafni A, Torrance G. Risk attitude and time preference in health. Management Science 1984;30(4):440-51.

47. Hyland ME. Quality of life of leg ulcer patients: questionnaire and preliminary findings. J Wound Care 1994;3(6):294-8.

48. Augustin M, Dieterle W, Zschocke I, Brill C, Trefzer D, Peschen M, et al. Development and validation of a disease-specific questionnaire on the quality of life of patients with chronic venous insufficiency. Vasa 1997 Nov;26(4):291-301.

49. Launois R, Reboul-Marty J, Henry B. Construction and validation of a quality of life questionnaire in chronic lower limb venous insufficiency (CIVIQ). Qual Life Res 1996 Dec;5(6):539-54.

50. Flett R, Harcourt B, Alpass F. Psychosocial aspects of chronic lower leg ulceration in the elderly. West J Nurs Res 1994 Apr;16(2):183-92.

51. Roe B, Cullum N, Hamer C. Patients' perceptions of chronic leg ulceration. In: Cullum N, Roe BH, editors. Leg Ulcers; Nursing Management. A Research Based Guide.London: Balliere Tindall; 1998. p. 125-34.

52. Bosanquet N. Costs of venous ulcers: from maintenance therapy to investment programmes. Phlebology 1992;Suppl.(1):44-6.

53. Stinnett A, Mullahy J. Net health benefits: a new framework for the analysis of uncertainty in cost-effectiveness analysis. Medical Decision Making 1998;18:S68-S80.

54. Bridges JF. Stated preference methods in health care evaluation: an emerging methodological paradigm in health economics. Appl Health Econ Health Policy 2003;2(4):213-24.

55. Moher D, Cook JC, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUORUM statement. Lancet 1999;354(b):1896-900.

56. Kikta MJ, Schuler JJ, Meyer JP, Durham JR, Eldrup-Jorgensen J, Schwarcz TH, et al. A prospective, randomized trial of Unna's boots versus hydroactive dressing in the treatment of venous stasis ulcers. J Vasc Surg 1988 Mar;7(3):478-83.

57. Ohlsson P, Larsson K, Lindholm C, Moller M. A cost-effectiveness study of leg ulcer treatment in primary care. Comparison of saline-gauze and hydrocolloid

treatment in a prospective, randomized study. Scand J Prim Health Care 1994 Dec;12(4):295-9.

58. Capillas R, Cabré V, Gil AM, Gaitano A, Torra JE. Comparación de la efectividad y coste de la cura en ambiente húmedo frente a la cura tradicional. Rev ROL Enf 2000;23(1):17-24.

59. Meaume S, Gemmen E. Cost-effectiveness of wound management in France: pressure ulcers and venous leg ulcers. J Wound Care 2002 Jun;11(6):219-24.

60. Augustin M, Siegel A, Heuser A, Vanscheidt W. Chronic leg ulcers: cost evaluation of two treatment strategies. Journal of Dermatological Treatment 1999;10:S21-S25.

61. Wayman J, Nirojogi V, Walker A, Sowinski A, Walker MA. The cost effectiveness of larval therapy in venous ulcers. Journal of Tissue Viability 2000;10(3):91-4.

62. Bosanquet N, Franks P, Moffatt C, Connolly M, Oldroyd M, Brown P, et al. Community leg ulcer clinics: cost-effectiveness. Health Trends 1993;25(4):146-8.

63. Kerstein MD, Gahtan V. Outcomes of venous ulcer care: results of a longitudinal study. Ostomy Wound Manage 2000 Jun;46(6):22-9.

64. Vickery L, Coe N, Pearson NJ. The impact of Somerset leg ulcer service developments. Journal of Clinical Excellence 2000;2:55-9.

65. Morrell CJ, Walters SJ, Dixon S, Collins KA, Brereton LM, Peters J, et al. Cost effectiveness of community leg ulcer clinics: randomised controlled trial. BMJ 1998 May 16;316(7143):1487-91.

66. Korn P, Patel ST, Heller JA, Deitch JS, Krishnasastry KV, Bush HL, et al. Why insurers should reimburse for compression stockings in patients with chronic venous stasis. J Vasc Surg 2002 May;35(5):950-7.

67. McColl E, JAcoby A, Thomas L, Soutter J, Thomas R, Harvey E, et al. Design and use of questionnaires: a review of best practice applicable to surveys of health service staff and patients. Health Technol Assess 2003;5(31):99-127.

68. de Vaus DA. Surveys in social research. 3rd ed. London: UCL Press Ltd; 1991.

69. Dillman DA. Mail and thelephone surveys: The total design method. New York: John Wiley & Sons, Inc; 1978.

70. Kind P. The EuroQol instrument: an index of health-related quality of life. Quality of Life and Pharmacoeconomics in Clinical Trials. Philadelphia: Lippincott-Raven; 1996.

71. Ware J, Jr., Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Med Care 1996 Mar;34(3):220-33.

72. Melzack R. The short-form McGill Pain Questionnaire. Pain 1987 Aug;30(2):191-7.

73. Iglesias C, Torgerson D. Does length of questionnaire matter? A randomised trial of response rates to a mailed questionnaire. J Health Serv Res Policy 2000 Oct;5(4):219-21.

74. Iglesias CP, Birks YF, Torgerson DJ. Improving the measurement of quality of life in older people: the York SF-12. QJM 2001 Dec;94(12):695-8.

75. Iglesias CP, Birks Y, Nelson EA, Scanlon E, Cullum NA. Quality of life of people with venous leg ulcers: a comparison of the discriminative and responsive characteristics of two generic and a disease specific instruments. Qual Life Res 2005 Sep;14(7):1705-18.

76. Birks YF, Porthouse J, Addie C, Loughney K, Saxon L, Baverstock M, et al. Randomized controlled trial of hip protectors among women living in the community. Osteoporos Int 2004 Sep;15(9):701-6.

77. Chalmers I. Comparing like with like: some historical milestones in the evolution of methods to create unbiased comparison groups in therapeutic experiments. Int J Epidemiol 2001 Oct;30(5):1156-64.

78. Cartwright A. Some experiments with factors that might affect the response of mothers to a postal questionnaire. Stat Med 1986 Nov;5(6):607-17.

79. Jacoby A. Possible factors affecting response to postal questionnaires: findings from a study of general practitioner services. J Public Health Med 1990;12(2):131-5.

80. Iglesias C, Nelson EA, Cullum NA, Torgerson DJ. VenUS I: a randomised controlled trial of two types of bandage for treating venous leg ulcers. Health Technol Assess 2004 Jul;8(29):iii, 1-iii105.

81. Nelson EA, Iglesias CP, Cullum N, Torgerson DJ. Randomized clinical trial of four-layer and short-stretch compression bandages for venous leg ulcers (VenUS I). Br J Surg 2004 Oct;91(10):1292-9.

82. Iglesias CP, Nelson EA, Cullum N, Torgerson DJ. Economic analysis of VenUS I, a randomized trial of two bandages for treating venous leg ulcers. Br J Surg 2004 Oct;91(10):1300-6.

83. Medical Research Council. MRC guidelines for good clinical practice in clinical trials. Mitcham, Surrey: Aldridge Print Group; 1998.

84. Altman DG. Practical statistics for medical research. Second ed. London: Chapman & Hall; 2002.

85. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. J Natl Cancer Inst 1959 Apr;22(4):719-48.

86. Peto R, Pike MC, Armitage P, Breslow NE, Cox DR, Howard SV, et al. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II. analysis and examples. Br J Cancer 1977 Jan;35(1):1-39.

87. Collett D. Modelling Suvival Data in Medical Research. Text in Statistical Science. London: Chapman & Hall; 1994.

88. Manca A, Hawkins N, Sculpher MJ. Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility. Health Econ 2005 May;14(5):487-96.

89. Lin DY, Feuer EJ, Etzioni R, Wax Y. Estimating medical costs from incomplete follow-up data. Biometrics 1997 Jun;53(2):419-34.

90. Briggs A, Sculpher M, Buxton M. Uncertainty in the economic evaluation of health care technologies: the role of sensitivity analysis. Health Econ 1994 Mar;3(2):95-104.

91. Briggs AH, Gray AM. Handling uncertainty when performing economic evaluation of healthcare interventions. Health Technol Assess 1999;3(2):1-134.

92. Van Hout VA, Ai MJ, Gordon GS, Rutten FF. Cost, effects, and C/E ratios alongside a clinical trial. Health Economics 1994;3(5):309-19.

93. Fenwick E, Byford S. A guide to cost-effectiveness acceptability curves. Br J Psychiatry 2005 Aug;187:106-8.

94. Roberts C. The implications of variation in outcome between health professionals for the design and analysis of randomized controlled trials. Stat Med 1999 Oct 15;18(19):2605-15.

95. Sculpher M, Claxton K, Akehurst R. It's just evaluation for decision-making: recent developments in, and challenges for, costs-effectiveness research. In: Smith PC, Ginnelly L, Sculpher M, editors. Health policy and economics: oportunities and challenges. 1st ed. Berkshire: Open University Press; 2005. p. 8-41.

96. Claxton K. The irrelevance of inference: a decision-making approach to the stochastic evaluation of health care technologies. Journal of Health Economics 1999;18:341-64.

97. Claxton K, Sculpher MJ, Drummond MF. A rational framework for decision making by the National Institute for Clinical Excellence. Lancet 2002;360:711-5.

98. Iglesias CP, Claxton K. Comprehensive Decision-Analytic Model and Bayesian Value-of-Information Analysis : Pentoxifylline in the Treatment of Chronic Venous Leg Ulcers. Pharmacoeconomics 2006;24(5):465-78.

99. Nelson EA, Cullum N, Jones J. Venous leg ulcers. Clin Evid 2002 Jun;(7):1806-17.

100. Belcaro G, Cesarone MR, Nicolaides AN, De Sanctis MT, Incandela L, Geroulakos G. Treatment of venous ulcers with pentoxifylline A 6-month randomized double-blind, placebo controlled trial. Angiology 2002;53(Suple):S45-S47.

101. De Sanctis MT, Belcaro G, Cesarone MR, Ippololito E, Nicolaides AN, Incandela L, et al. Treatment of venous ulcers with pentoxifylline A 12-month double-blind, placebo controlled trial. Microcirculation and healing. Angiology 2002;53(Suple):S49-S51.

102. Cooper NJ, Sutton AJ, Abrahms KR, Turner D, Wailoo A. Comprehensive decision analytical modelling in economic evaluation: A Bayesian approach. Health Economics 2000;13:203-26.

103. Whitehead A, Whitehead J. A general parametric approach to the meta-analysis of randomized clinical trials. Stat Med 1991 Nov;10(11):1665-77.

104. Spiegelhalter DJ, Thomas A, Best NG, Lunn D. WinBUGS Version 1.4 User manual. Available at www.mrc-bsu.cam.ac.uk/bugs. Cambridge: MRC Biostatistics Unit; 2002.

105. Schurmann W, Eberhardt R. The efficacy of pentoxifylline added to topical and compression therapy in patients with varicose and postthrombotic leg ulcers. Therapiewoche 1986;36:2343-5.

106. Colgan MP, Dormandy J, Jones P, Schraibman I, Shanik G, Young R. Oxpentifylline treatment of venous ulcers of the leg. British Medical Journal 1990;300:972-5.

107. Falanga V, Fujitani RM, Diaz C, Hunter G, Jorizzo J, Lawrence PF, et al. Systemic treatment of venous leg ulcers with high doses of pentoxifylline: efficacy in a randomized, placebo-controlled trial. Wound Repair and Regeneration 1999;7(208):213.

108. Claxton K, Neumann PJ, Araki S, Weinstein MC. Bayesian value-of-information analysis. An application to a policy model of Alzheimer's disease. Int J Technol Assess Health Care 2001;17(1):38-55.

109. UK National Helath Service Health technology Assessment Programme. http://www.ncchta.org/ProjectData/3_project_record_published.asp?PjtId=1076&SearchText=leg+ulcers [Accessed May 2005].

110. Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. Stat Med 2004 Oct 30;23(20):3105-24.

111. Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. BMJ 2005 Oct 15;331(7521):897-900.

112. Ades AE, Claxton K, Sculpher M. Evidence synthesis, parameter correlation and probabilistic sensitivity analysis. Health Econ 2006 Apr;15(4):373-81.

113. Sutton AJ, Abrams KR, Jones DR. Generalized synthesis of evidence and the threat of dissemination bias. the example of electronic fetal heart rate monitoring (EFM). J Clin Epidemiol 2002 Oct;55(10):1013-24.

114. Sutton AJ, Abrams KR. Bayesian methods in meta-analysis and evidence synthesis. Stat Methods Med Res 2001 Aug;10(4):277-303.

115. Edwards P, Roberts I, Sandercock P, Frost C. Follow-up by mail in clinical trials: does questionnaire length matter? Control Clin Trials 2004 Feb;25(1):31-52.

116. Perneger TV, Burnand B. A simple imputation algorithm reduced missing data in SF-12 health surveys. J Clin Epidemiol 2005 Feb;58(2):142-9.

117. VenUS II research team. Venus II: larval therapy venous ulcer study www.ncchta.org/projectdata/1_search_list1.asp [Accessed June 2006].

118. VenUS III research team. VenUS III - Venous Ulcer Studies III: ultrasound for venous leg ulcers. http://www.ncchta.org/projectdata/1_search_list1.asp [Accessed June 20006].

119. Lamping DL, Schroter S, Kurz X, Kahn SR, Abenhaim L. Evaluation of outcomes in chronic venous disorders of the leg: development of a scientifically rigorous, patient-reported measure of symptoms and quality of life. J Vasc Surg 2003 Feb;37(2):410-9.

120. Brazier JE, Roberts J. The estimation of a preference-based measure of health from the SF-12. Med Care 2004 Sep;42(9):851-9.

121. Therneau TM GP. Frailty Models. In: Therneau TM GP, editor. Modeling Survival Data: Extending the Cox Model.Heidelberg: Springer Verlag; 2000.

122. Torra JE, Rueda J, Blanco J, Torres J, Toda L. Úlceras venosas ¿Sistema de compresión multicapa o venda de crepé? Estudio comparativo sobre la efectiivdad, coste e impacto en la calidad de vida. Rev ROL Enf 2003;26(6):471-8.

123. Carr L, Phillips Z, Postnett J. Evaluation of the relative cost-effectiveness of profore. York Health Ecnomics Consortium; 1998.

124. Taylor AD, Taylor RJ, Marcuson RW. Prospective comparison of healing rates and therapy costs for conventional and four-layer high-compression bandaging treatments of venous leg ulcers. Phebology 1998;13(1):20-4.

125. Schonfeld WH, Villa KF, Fastenau JM, Mazonson PD, Falanga V. An economic assessment of Apligraf (Graftskin) for the treatment of hard-to-heal venous leg ulcers. Wound Repair Regen 2000 Jul;8(4):251-7.

126. Ades AE, Sculpher M, Sutton A, Abrams K, Cooper N, Welton N, et al. Bayesian methods for evidence synthesis in cost-effectiveness analysis. Pharmacoeconomics 2006;24(1):1-19.

127. Spiegelhalter DJ, Abrams KR, Myles JP. Bayesian approaches to clinical trials and health-care evaluation. The Atrium, Southern Gate, Chichester: John Wiley & Sons Ltd.; 2004.

128. Margolis DJ, Berlin JA, Strom BL. Which venous leg ulcers will heal with limb compression bandages? Am J Med 2000 Jul;109(1):15-9.

129. Margolis DJ, Berlin JA, Strom BL. Risk factors associated with the failure of a venous leg ulcer to heal. Arch Dermatol 1999 Aug;135(8):920-6.

130. Margolis DJ, Berlin JA, Strom BL. A sensitivity analysis of risk factors predicting success treating a venous leg ulcer. Wound Repair and Regeneration 1998;A246 Abstract.

131. Danielsen L, Madsen SM, Henriksen L. Venous leg ulcer healing: a randomized prospective study of long stretch versus short stretch compression bandages. Phlebology 1998;13:59-63.

132. Pocock SJ, Simon R. Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. Biometrics 1975 Mar;31(1):103-15.

133. Bradley M, Cullum N, Sheldon T. The debridement of chronic wounds: a systematic review. Health Technol Assess 1999;3(17 Pt 1):1-78.

134. Hansson C, Persson L-M, Stenquist B, Nordin P, Roed-Petersen J, Westerhof W, et al. The effects of cadexomer iodine paste in the treatment of venous leg ulcers compared with hydrocolloid dressing and paraffin gauze dressing. International Journal of Dermatology 1998;37(5):390-6.

135. Smith JM, Doré CJ, Charlett A, Lewis JD. A randomized trial of biofilm dressing for venous leg ulcers. Phlebology 1992;7:108-13.

136. Barber JA, Thompson SG. Analysis of cost data in randomized trials: an application of the non-parametric bootstrap. Stat Med 2000 Dec 15;19(23):3219-36.

137. Thompson SG, Barber JA. How should cost data in pragmatic randomised trials be analysed? BMJ 2000 Apr 29;320(7243):1197-200.