Is acupuncture an effective treatment for menorrhagia?

Systematic review, exploratory randomised trial, qualitative investigation and GP survey

Volume 1.

Alison Elizabeth Longridge

BA DipAc MScAc MBAcC

Submitted in partial fulfilment of the requirements for a PhD Degree

Department of Health Sciences

University of York

January 2006

Abstract

The systematic review in Chapter Two reveals little compelling evidence for acupuncture in the treatment of menorrhagia and other menstrual disorders. The York and Selby Acupuncture for Menorrhagia Exploratory Trial (ACUMEN) was undertaken to explore the potential for a definitive pragmatic trial to evaluate the relative clinical and cost effectiveness of acupuncture in primary care.

ACUMEN was a rigorous pragmatic randomised trial set within Selby and York Primary Care Trust. Nineteen patients randomised to the comparator arm received usual care from their GPs. Twenty patients randomised to the experimental arm were also *offered* referral for traditional acupuncture to one of three professional acupuncture practitioners in York. Despite a total sample size of only 39, the findings are promising. Chapter Three concludes that a full trial is justified, but highlights recruitment as an issue

The longitudinal qualitative enquiry nested in ACUMEN and reported in Chapter Four provides evidence of the acceptability of acupuncture treatment and how this evolved over time for six women in ACUMEN. In spite of its small sample size, its methodological rigour and degree of congruence with previous work generate insight into: why patient-centred recruitment strategies were successful in ACUMEN; why patients complied fully with acupuncture; and the nature of the complex package of 'active ingredients' acupuncture patients experienced within ACUMEN. Moreover these findings offer guidance for the design of recruitment material and the delivery of acupuncture in future trials.

Chapter Five reports a complementary postal survey to investigate barriers to patient recruitment by GPs. This leads to recommendations for procedures to overcome these barriers and guard against failure to complete a full-scale trial. Chapter Six considers the evidence from ACUMEN in the light of past research, concludes that a full trial is justified and feasible, and highlights design issues.

LIST OF CONTENTS

Chapter Three: The York and Selby ACUMEN Exploratory Trial

LIST OF APPENDICES

LIST OF TABLES

Chapter Three: The York and Selby ACUMEN Exploratory Trial

Chapter Four: The ACUMEN Qualitative Study

Chapter Five: The ACUMEN GP Survey

LIST OF FIGURES

Preface

I was one of the first students to study for a Masters in Acupuncture at the Northern College of Acupuncture in York and qualified with a distinction in May 2000, almost six years ago. This primarily literature based study with six complementary case studies was an exploration of the treatment of primary dysmenorrhoea by traditional acupuncture. This study furthered my interest in the potential role for acupuncture in the treatment of menstrual disorders and led me to adopt this as my special interest in clinical practice. It also led to my writing a briefing paper for the Acupuncture Research Resource Centre on the evidence of effectiveness for acupuncture in the treatment of gynaecological conditions. This enabled me to understand better what makes a good effectiveness study and whether such research is possible for a holistic medical paradigm like traditional Chinese medical acupuncture. A prominent concern within the profession at that time was that, whilst qualitative research affords the methods to investigate patients' treatment experiences and perceptions of benefits for example, quantitative research threatens only to divorce acupuncture from its theoretical principles, transforming it utterly by reducing it to a rigid set of points uniformly administered within the confines of the laboratory style experimental model. Whilst the York Acupuncture for Back Pain (YACBAC) trial was raising awareness about rigorous studies of individualised acupuncture as practiced in 'the real world' by members of the British Acupuncture Council, there were concerns that such studies would ultimately be thwarted and acupuncture's effect systematically underestimated owing to a lack of tools to detect and estimate accurately the range and type of benefits that patients experience and value.

The studentship at the Department of Health Sciences provided me with a unique opportunity to begin to grapple with some of these questions. I also took this opportunity to move from the study of primary dysmenorrhoea to that of menorrhagia, as it was a condition which I had had some success treating, that caused women much distress, and which was *not* characterised by pain – a characterisation of acupuncture that I was keen to leave behind. But as I was the sole acupuncturist in the department, it was a journey that was not without its challenges. It took me eighteen months to reach the point where I felt able to assert my research agenda and approach those with the expertise and open mindedness that I needed to help me achieve my aim. I am indebted to Professors Ian Russell and Kate Thomas for accepting my request to supervise my studies and to Dr Hugh MacPherson for providing financial backing for a trial in his role as Director of the Foundation for Traditional Chinese Medicine. Within months ACUMEN had come into being and epitomised the wish *"to do justice to holism"* – words taken from a colleague's MSc thesis (Alison Gould), which I wrote on my application for a studentship at the Department of Health Sciences, University of York.

Acknowledgements

I am indebted to the following people for their guidance, support and encouragement during the research: Professor Ian Russell for his contribution to the design of the entire programme of research and the conduct and analysis of the systematic review, exploratory trial and postal questionnaire survey specifically; Professor Kate Thomas for her contribution to the qualitative investigation and help with methodological aspects of the exploratory trial and survey; Dr Hugh MacPherson for his help with methodological aspects of the exploratory trial and survey; Professor Christine Godfrey for her contribution to the economic protocol and help with other methodological aspects of the exploratory trial and survey; Dr Daphne Russell for her statistical advice on the trial analysis and help with the SF-36 and EQ-5D syntax; Dr Manuela Fontebasso for her help with methodological aspects of the exploratory trial and survey; Janette Colclough for her help with methodological aspects of the systematic review; Ms Fenella Jeffers, Ms Alison Gould and Ms June Tranmer for their help with methodological aspects of the exploratory trial; the thirty-nine women who participated in the ACUMEN trial, with especial thanks to those not randomised to the option of acupuncture and those who participated in interviews; Dr Fiona Fylan for her guidance in the final stages of the thesis; all my colleagues at both the Department of Health Sciences at the University of York and Northern College of Acupuncture in York for their feedback, ideas, encouragement and humour; my colleagues and patients at the York Clinic for Complementary Medicine for their encouragement and continuing source of inspiration; and above all my family for their continued and unfaltering support.

I am grateful to the Department of Health Sciences for awarding me a research studentship. I am also grateful to the Foundation for Traditional Chinese Medicine for providing office premises, meeting the administrative costs of the research, and funding the acupuncture treatments for the trial patients.

Authors Declaration

The candidate is solely responsible for all the research presented in this thesis. She conducted the entire review with assistance from the departmental librarian and the second rater. Whilst her multi-disciplinary research advisory group provided invaluable guidance, the candidate undertook the design, implementation, analysis and reporting of the trial, the qualitative investigation and the survey.

Chapter One

Is acupuncture an effective treatment for menorrhagia?

An Introduction to the Thesis

1. Background

1.1 The nature of menorrhagia

Menorrhagia is defined as excessive or prolonged menstrual bleeding in the absence of discernable organic pathology, such as fibroids or endometriosis (Shaw et al. 2002). It can result in physical, psychological and social impairment (Lumsden & Smith 1992; Vessey et al. 1992), and is likely to have important effects on employment, though no studies have assessed the extent of this problem. When objectively assessed, what constitutes 'excessive' or 'prolonged' is a measured menstrual blood loss in excess of 80 ml, which is thought to affect approximately 10% of women of reproductive age. The population studies from which these figures are derived defined 80 ml as the upper limit for normal menstrual blood loss on the basis that it is women with a menstrual blood loss in excess of 80 ml whom are at greater risk of developing iron deficiency anaemia (Barer & Fowler 1936; Halberg et al. 1966; Cole et al. 1971), and menorrhagia is the commonest cause of iron deficiency anaemia in the developed world (Cohen & Gibor 1980).

What these studies also revealed, however, was that this objective measure does not always correlate with women's subjective assessments of their menstrual loss. Halberg and colleagues (1966) found 40% of the women who met the criteria for objectively defined menorrhagia to consider their loss moderate or light, whilst 26% of those with a measured loss of 60 ml or less described their menses as heavy. Similarly, the studies by both Fraser and colleagues (1984) and Cameron and colleagues (1990) found 38% and 50%, respectively, of women presenting with menorrhagia to have a measured loss within the normal range. It is notable, therefore, that whilst objectively assessed menorrhagia is estimated to affect 10% of women of reproductive age, in a national survey of women conducted in the UK in 1990, 30% - almost a third of all women - reported that they experienced symptoms of heavy

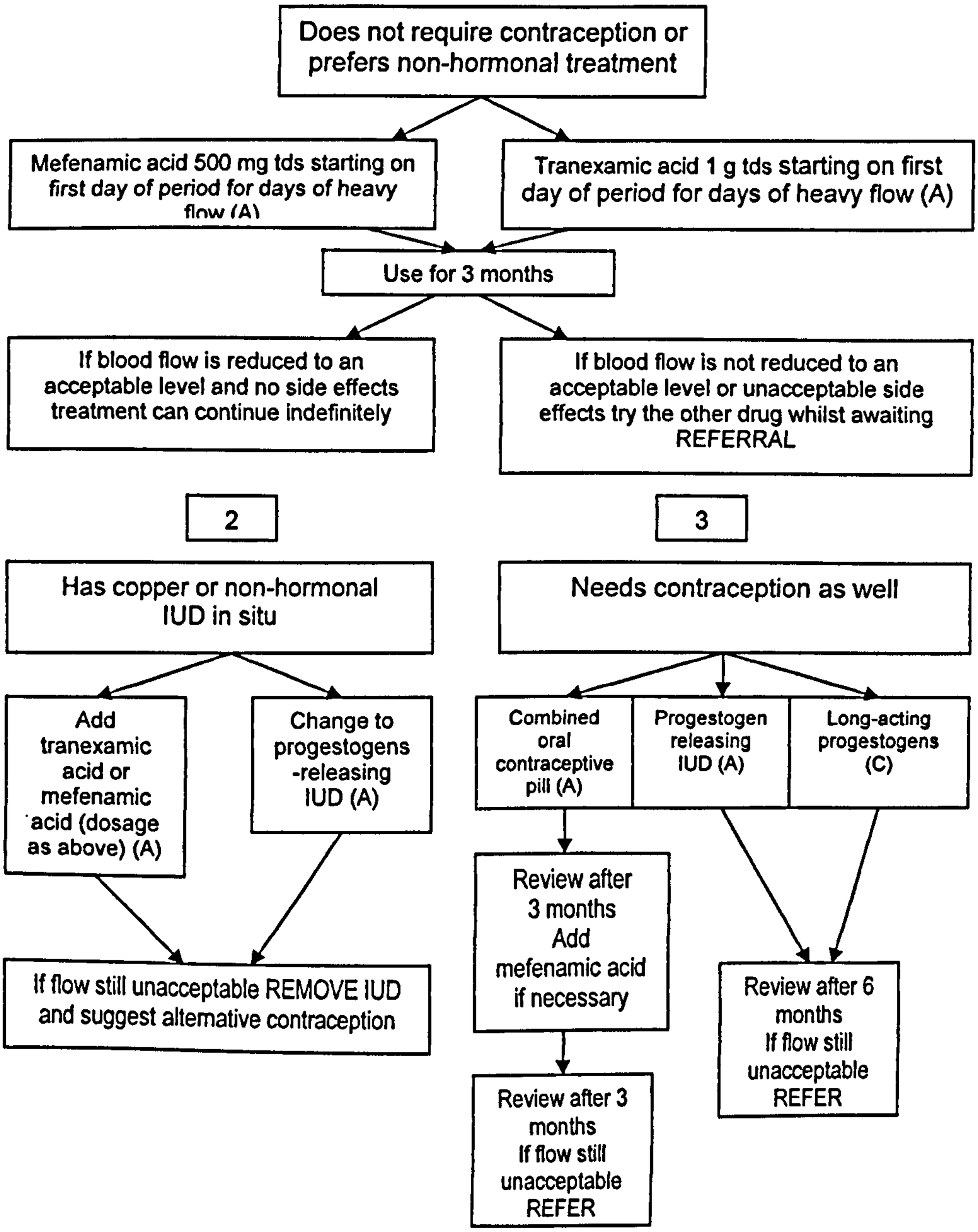
menstrual loss (MORI 1990). This finding is helpful, as it is more likely to reflect the extent of the demand for medical care for menorrhagia.

A number of explanations have been given to explain the discrepancy between measured and perceived menstrual loss. Fraser and colleagues (1985) argued that the alkaline haematin method used to provide an objective measure of menstrual loss may be inadequate, because of the considerable variation in blood to fluid ratio between women, whereby the presence or absence of anaemia is not predictive of the volume of fluid loss and the distress this may cause. However, the most widely reported explanation is that because women have no perception of what their actual menstrual loss is, they can only make a subjective assessment based on their own experience or that of other women (Rees 1991; Miles 1991; Prentice 1999a), and psychological factors are also thought to influence their perception and tolerance of symptoms (Heiman 1956; Greenberg 1983; Marshall 1998). Thus, clinicians and researchers have suggested that counselling, advice and reassurance, may be more appropriate than any active intervention in women whose measured menstrual blood loss is within the normal range (Rees 1991; Coulter 1995; Marshall 1998), especially given the ethical implications of prescribing them medication that carries the risk of unpleasant, and potentially serious side effects (Lumsden & Smith 1992; Coulter et al. 1995; Irvine & Cameron 1999). The problem with recommending that management be based on such laboratory criteria is twofold: First, the alkaline haematin method is a specialised and time-consuming technique that is not available for routine clinical use. And second, it is a technique that requires women to collect their used sanitary towels and tampons, a process which many women find unacceptable. Thus, general practitioners (GPs) have continued to employ a more pragmatic approach, whereby a woman's perception and experience of menstrual loss, and the degree to which symptoms adversely affect quality of life, play a key role in the therapeutic decision making process.

1.2 The medical management of menorrhagia

Menorrhagia is a common reason for GP consultation (Nuffield Institute & Centre for Reviews and Dissemination 1995), with 5% of women in the UK seeking help for this symptom annually (Vessey et al. 1992). It is also responsible for 12% of all referrals

Figure 1: **Medical management of the complaint menorrhagia** from *"Guideline: The Initial Management of Menorrhagia"* published by the Royal College of Obstetrics and Gynaecology, UK www.rcog.org.uk/guidelines/menorrhagia.html 01/02/00



Whilst oral luteal phase progestogens are ineffective in reducing menstrual blood loss (A), intrauterine progesterones are effective (A)

A = based on randomised trials
B = based on other robust experimental or observational studies
C = based on more limited evidence but the advice relies on expert opinion and has the endorsement of respected authorities

Note: This guideline was in date at the time of the ACUMEN Study

to gynaecology outpatient departments (Bradlow et al 1992), and at least 60% of those referred for menorrhagia will undergo surgical intervention (Coulter et al 1989; Coulter et al 1991). About half of all women who undergo surgery for menorrhagia have a normal uterus (RCOG 1998). So it is important to provide women with less invasive and more supportive treatment options for menorrhagia.

The treatment objective in the initial management of menorrhagia is to alleviate heavy menstrual flow and, as a consequence, prevent iron deficiency anaemia and improve quality of life. The drug therapies available fall into four main categories: (1) Hormonal treatments (e.g. norethisterone, danazol and hormone releasing coils); (2) prostaglandin synthetase inhibitors (e.g. mefenamic acid and naproxen); (3) inhibitors of fibrinolysis (e.g. tranexamic acid); (4) and reducers of platelet fragility (e.g. ethamsylate). In 1995, Coulter and colleagues conducted a systematic review of randomised controlled trials of drug therapies for menorrhagia, which later informed the development of the national guideline for the initial management of menorrhagia published by the Royal College of Obstetrics and Gynaecology in 1998 *[Figure 1]*. Coulter concluded that tranexamic acid, danazol and progestagen-releasing coils (Mirena coil) could significantly reduce both objective and subjectively assessed menstrual blood loss. The Mirena coil has not yet been licensed for use in Britain, but it is thought it may provide an effective alternative to surgery and be useful in patients who also require contraception (Irvine et al 1997; Prentice 1999a, b). The considerable risk of becoming amenorrhoeic after 12 months use may, however, make it an unacceptable option for many women. Danazol is also extremely effective at reducing menstrual blood loss, but its use is restricted due to the considerable risk of developing serious androgenic side effects (Coulter et al 1995; Irvine et al 1997). Thus, the most effective and acceptable of the drug therapies currently available is tranexamic acid, which induces a reduction in menstrual blood loss of around 50% in the majority of patients (Coulter et al 1995; Bonnar & Sheppard 1996). However, tranexamic acid is ineffective in around 15% of patients, and its' acceptability is compromised in approximately 30% of patients, owing to the frequency of side effects such as headache, skin rashes and gastrointestinal symptoms (Nilsson & Rybo 1967). It is also an inappropriate drug therapy for women with thromboembolic disease and coagulation or fibrinolytic disorders (BNF 2005), and a treatment that achieves a reduction in blood loss, at the expense of unpleasant side effects is unlikely

to be acceptable when considering quality of life terms (Coulter et al 1995; Irvine et al 1999; Lumsden & Smith 1992). Moreover, studies have found that many women find long-term drug therapy unacceptable (Marshall 1998; Coulter et al 1995), especially if their symptoms will return on cessation of medication. Unless they can be reassured on this point, they may find surgery a more attractive option. A preference for treatment that does not involve drugs or surgery has also been reported by women (Longridge nee Gamon 2000; Scambler & Scambler 1985).

The premise of this thesis is that acupuncture might usefully add to the therapeutic options currently available to women with menorrhagia in NHS primary care, so enabling more women to manage their symptoms without recourse to surgery. In particular, acupuncture might better address the health needs of those women who have found drug therapy ineffective, inappropriate because of side effects or concurrent medical conditions, and unacceptable because they dislike drug therapies.

1.3 Acupuncture in Britain today

Within Britain, practitioners of acupuncture can be broadly defined as practicing either traditional acupuncture or western medical acupuncture (Acupuncture Regulatory Working Group 2003). Traditional acupuncture involves completing a comprehensive three-year undergraduate training programme based on oriental medical principles originating in China, and the study of western medical sciences appropriate to the practice of acupuncture and independent practice in healthcare. Most traditional acupuncturists use acupuncture as their main or only form of treatment. Western medical acupuncture is a system of practice that focuses on those aspects of acupuncture which may be easily integrated into orthodox medical practice, or which are amenable to physiological explanation and evaluation through laboratory tests or explanatory trials (Filshie et al. 1998; James 1998). A basic competence course in medical acupuncture typically involves two weekends, a safety assessment and the submission of 30 case reports. It is practiced by doctors, physiotherapists, nurses and dentists and is most often regarded as an adjunct to existing treatments.

This thesis is primarily concerned with traditional acupuncture, as practiced by Members of the British Acupuncture Council (MBAcC). The British Acupuncture

Council was formed in 1995 and is a non-statutory regulatory body (BAcC). The BAcC maintains standards of education, ethics, discipline and practice to ensure the health and safety of the public at all times. Members are bound by the Council's Codes of Safe Practice, Professional Conduct and Health & Disciplinary Procedures, and are covered by full Medical Malpractice & Public/Products Liability Insurance. To achieve BAcC membership, practitioners must first attain a recognised qualification. In Britain, BAcC recommended training institutions are those accredited by the British Acupuncture Accreditation Board (BAAB), an independent body closely allied to the BAcC that was established in 1990 (The British Acupuncture Council 2004). To foster the development of a research culture within

the profession, the BAcC has stipulated that research methods be taught at colleges as part of the core curriculum. The Acupuncture Resource Research Centre (ARRC) was set up in 1994 by the BAcC to facilitate undergraduate and postgraduate studies better. In 2004 there were eight fully accredited training institutions and the BAcC had a membership of over 2500 (The British Acupuncture Council 2004).

1.4 Acupuncture in the treatment of gynaecological conditions – past and present

Gynaecology in Chinese medicine has a long history dating back to 1500 – 1000 BC (Maciocia1997). *The Yellow Emperor's Classic of Internal Medicine: Simple Questions (1979)* was first published 100 BC, and has many references to women's physiology and anatomy, including a description of the aetiology and treatment of menorrhagia and dysmenorrhoea. From 206 BC, texts dealing exclusively with gynaecology were written. Over the course of two millennia, through a continuous process of critical thinking and extensive clinical observation and testing, the theory and practice of acupuncture for gynaecological diseases has been developed. It is a process and a history that since the 1960's, the West has participated in (Birch & Felt 1999a). Seminal texts such as Maciocia's *Obstetrics & Gynecology in Chinese Medicine* (1997) continue the tradition of balancing old and new in order to fulfil the mandate of accurate transmission and appropriate modulation, be it from one generation or one country to the next. Classical citations are balanced with contemporary case histories and diseases common in Western societies, but rare in China, are innovatively presented (Kaptchuk in Maciocia 1997).

Table 1: Illustration of the range of variables significant to the process of diagnosis and treatment in TCM – *from a case report of "Lydia" (Longridge nee Gamon 2000)*

Pre-menstrual symptoms:
- Feeling "wound-up"
- Irritability
- Tiredness
- Abdominal bloating
- Menstrual symptoms:
- Abdomen feels heavy and dragging
- Low backache more pronounced
- Heavy ache down front of thighs
- Severe cramping abdominal pain that is fixed
- Pain stops as soon as clots are passed
- Heavy, dark, clotted loss days 1 and 2
- Clots size of 50p
- Bleed brighter in colour and lighter flow days 3 to 6

Other symptoms and signs:
- Abdominal bloating after eating
- Low back ache – constant ache and weakness
- Loose, strong smelling stools Dark, strong smelling urine
- Feels the heat and sweats easily
- Quite overweight since childhood, family pattern
- Malar flush Pulse quality: slippery, wiry and rapid
- Tongue: pink, wet, redness and swelling in central area, many heat spots, thick greasy yellow coat at the root

Other factors:
- No breakfast as appetite poor am.
- Irregular diet due to work environment (hospital nurse)
- Snacks on convenience foods (crisps, chocolate, sandwiches)
- Main meal often a take-away (Chinese, Indian, fish and chips)
- 4 to 6 cups of coffee per day
- 1 to 2 bottles of wine one evening each week
- Lots of cold water throughout the day
- Too tired to exercise
- Frustration and worry related to financial problems and work
- Good personal and working relationships
- Supportive family close by

Table 2: Illustration of a TCM differential diagnosis according to syndrome patterns – *from a case report of "Lydia" (Longridge nee Gamon 2000)*

Spleen Qi Xu and Damp
- Tiredness
- Tendency to be overweight
- Abdominal bloating
- Appetite poor am
- Pulse: slippery
- Tongue: wet, swelling in central area

Liver Qi Stagnation
- Feeling "wound-up"
- Irritability
- Pulse: wirey
- Damp Heat in the Lower Jiao
- Abdomen feels heavy and dragging during menses
- Low back ache during menses
- Heavy ached down front of thighs during menses
- Loose, strong smelling stools
- Dark, strong smelling urine
- Feels the heat and sweats easily
- Drinks a lot of cold water throughout the day
- Pulse: slippery and rapid
- Tongue: redness and swelling in central area, many heat spots, thick greasy yellow coat at the root

Kidney Yin Xu with Empty Heat
- Low back ache – constant dull ache, weakness
- Malar flush
- Feels the heat and sweats easily
- Pulse: rapid
- Stagnation of Blood in the Uterus and Chong Mai
- Severe cramping abdominal pain that is fixed

Today, traditionally trained acupuncturists in the West are taught to diagnose and treat a range of gynaecological disorders including menorrhagia, dysmenorrhoea, and irregular menstruation, according to oriental medical principles originating in China (Flaws 1997; Kaptchuk 1983; Maciocia 1989; Maciocia 1997; Maciocia 1998). This holistic medical paradigm requires that the patient's specific and general physiological and psychological response to a disease entity is assessed by means of questioning, observation, palpation, listening and smelling *[Table 1]*. The relationships between the patient's signs and symptoms are then categorised and a differential diagnosis made *[Table 2]*. The context in which the disease takes place, the patient's living and working environment, is also taken into consideration in order to understand the disease process and create the most effective treatment strategy *[Figure 2 & Table 3]*. This approach tends to the unique way in which disease manifests itself in individuals and involves the tailoring of treatment to the individual.

The emphasis is also on education and self-care according to aetiology, whereby advice on diet, rest and exercise is offered as appropriate. Consequently the quality of the patient-practitioner relationship is understood to be an important factor in bringing about health and healing (MacPherson & Kaptchuk 1997).

Table 3: Illustration of a treatment protocol – *from a case report of "Lydia" (Longridge nee Gamon 2000)*

Points, point actions and needling details

- **SP6** sanyinjiao, strengthens the Spleen, resolves Damp, moves the Blood and Nourishes Kidney Yin - bilateral, even method, deqi obtained [needle type: Seirin 0.20mm x 25mm]
- **ST36** taichong, regulates Liver Qi and moves Blood – bilateral, reducing method, deqi obtained [needle type: Seirin 0.20mm x 25mm]
- **REN3** zhongji, resolve Dampness and Blood Stagnation in the Uterus – reducing method, deqi obtained [needle type: Seirin 0.20mm x 25mm]
- **ST28** shuidao, resolves Dampness and Blood Stagnation in the Uterurs – reducing method, deqi obtained [needle type: Seirin 0.20mm x 25mm]
- **SP4** gongsun and **P6** neiguan, regulate menstruation – right-left crossover, even method, deqi obtained [needle type: Seirin 0.16mm x 25mm]

Chinese Dietary Advice to strengthen the Spleen and resolve Damp

Avoid: dairy (goat and sheeps products are easier to digest), wheat, sugar, yeast, bananas, fruit juices from concentrate especially orange and tomato, fried foods, peanuts, bananas, pork and rich meat

Include: anchovy, mackerel, aduki beans, garlic, onion, celery, parsley, olive, pear, grapefruit, watercress, pumpkin, dark leafy green vegetables, brown rice, dried apricot and fig

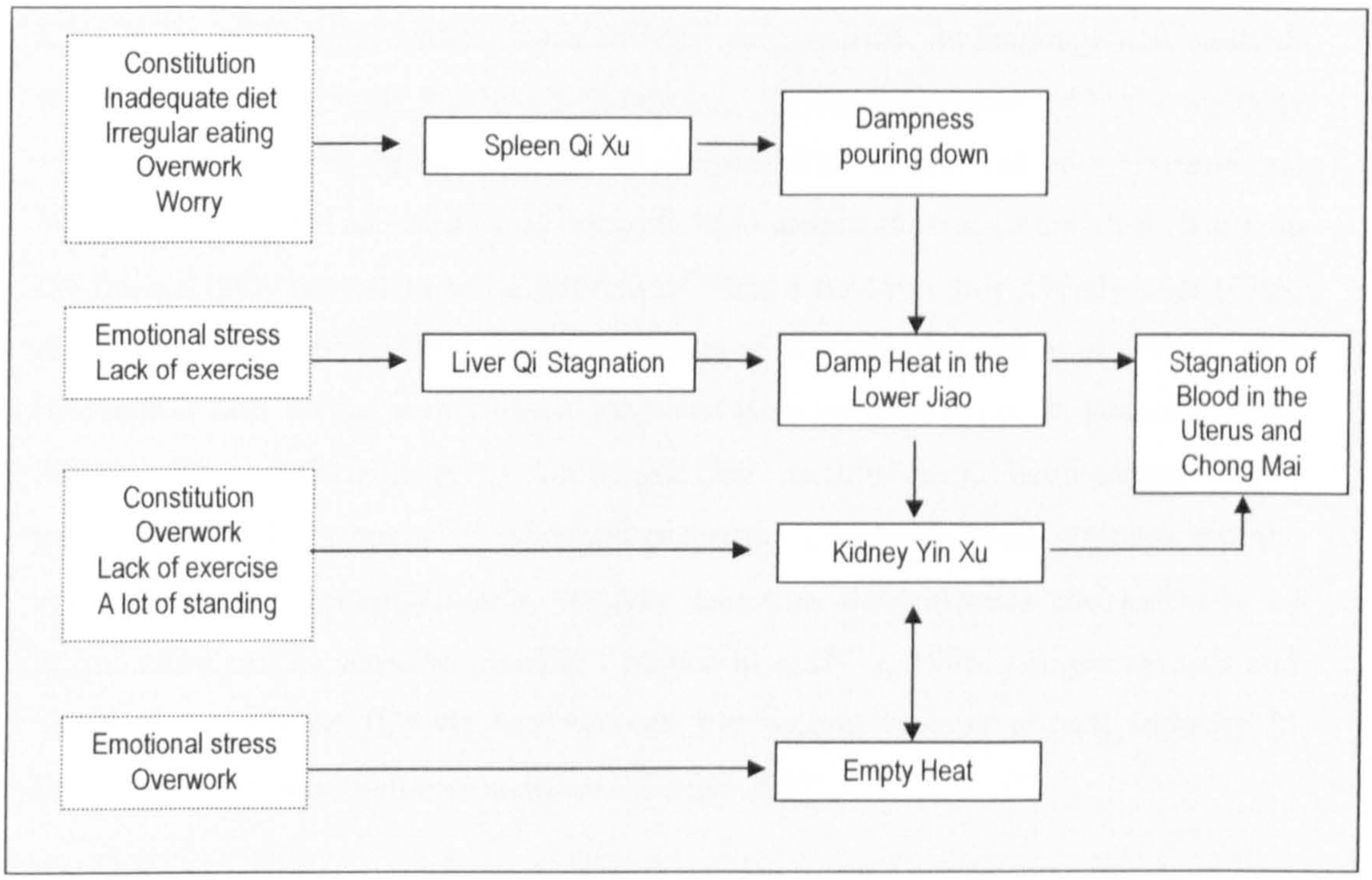
Qi Gong – to tonify Kidney Yin

Aerobic activity – walking or cycling to regulate Liver Qi, move Blood and resolve Dampness

Relaxation – to nourish Kidney Yin, strengthen the Spleen and soothe the Liver

It is generally accepted that treatment for menstrual problems is required over the course of a minimum of three menstrual cycles, involving between 12 and 20 sessions (Lyttleton 2004; Maciocia 1998). The treatment itself involves an individualised acupuncture-points prescription and the insertion of fine (approx. 0.20 gauge), sterilised, disposable needles into the skin and underlying tissues at precise points on the body (British Acupuncture Council Code of Safe Practice Committee 2002; Deadman et al. 1998). Between six and 12 local and distal points are commonly used. The needle technique, that is the way in which the point is manually stimulated or the use of electricity to stimulate the point (electro-acupuncture), is determined by the treatment principle *[Table 3]* (MacPherson & Kaptchuk 1997). Other interventions such as moxibustion, infra-red heat lamp, massage, and cupping are administered as indicated (Birch & Felt 1999b). Moxibustion involves igniting the processed mugwort plant "moxa" so that the heat from the burning herb stimulates the acupoint or an area of the body, such as the abdomen. Infrared heat lamps can also be used to heat areas of the body. Massage or cupping, the application of suction cups to the body, can be used to stimulate circulation. The precise nature of the intervention may change over time in line with the patient's health state and response to treatment.

Figure 2: Illustration of a Pathology and Aetiology Diagram - *from a case report of Lydia (Longridge nee Gamon 2000)*



Traditional acupuncture has its roots in Daoist philosophy (Birch & Felt 1999c; Kaptchuk 1983; Matsumoto & Birch 1983), and it is this theoretical underpinning that allows for the holistic treatment approach described. In Daoist thought the cosmos, heaven and earth, is perceived as an integral whole that gives rise to phenomenon. Creation is seen as a continual process that occurs by a kind of inductance, a movement that comes about as a result of an inner dynamic of cyclical patterns that arise as a result of the existence of two polar complements or forces labelled Yin Qi and Yang Qi. Human beings are seen as being connected to this all-encompassing energetic system, and to be microcosms of the macrocosm. Mind and body are seen as inter-connected and the dynamic flux and cyclical pattern of change observed in nature is also observed in human beings. As a consequence, health is seen not as an objective state, but as a dynamic harmony. Disease is seen as a way of communicating that the dynamic flux between ourselves and the world we live in has broken down at some level, be it mind (i.e. psychological), body (i.e. physiological) or spirit (i.e. the way we infuse purpose or meaning into our lives). The principle of

treatment is to help re-establish this ebb and flow of Qi and in so doing generate the best possible internal climate for self-repair and recovery.

Clearly the philosophy underlying traditional acupuncture, its language and methods differ greatly to those of Western biomedicine. These differences led to the majority of Western scientists being sceptical of acupuncture. It also led both Chinese and Western biomedical scientists to investigate how acupuncture accomplishes change in the human body according to the models of human function they already trust (Chen & Yu 1991; Han 1997; Mo et al. 1993; Omura et al. 1998; Takeshi et al. 1976). This research is still in the very earliest stages of development (Stux & Hammerschlag 2001). Thus, left entirely to the biomedical establishment, acupuncture would probably never have begun its westward progress. Unschuld (1992) suggests that the current extent of acupuncture's Western adoption demonstrates the rationale of acupuncture makes sense to patients. Indeed as early as 1986, complementary and alternative medicine (CAM) had become the second biggest growth industry in Europe, second only to microelectronics (Reilly 2001).

CAM was best defined by the Cochrane Collaboration for the House of Lords Report on Complementary and Alternative Medicine (2000) as "*a broad domain of healing resources that encompasses all health systems, modalities, and practices and their accompanying theories and beliefs, other than those intrinsic to the politically dominant health systems of a particular society or culture in a given historical period*". Perhaps what is most striking about the rise in CAM usage is that whereas most of the great revolutions in medicine have conceivably come from scientists, doctors, or technology, this phenomenon has been patient-led – a publicly driven trend (Reilly 2001). Approximately 7% of the adult population in England have received acupuncture, the majority of whom are women (Thomas et al. 2001). Professional acupuncturists, of whom there are over 2,500 in the UK (The British Acupuncture Council 2004), provide the great majority of these treatments, and at least 6% of those patients consulting professional acupuncture practitioners in the UK receive treatment for gynaecological and obstetric conditions (Newton et al. 2002; Wadlow & Peringer 1996).

1.5 Acupuncture research

Indeed, it is the rapidly increasing number of patients who are seeking help from CAM practitioners that is the main reason for the current upsurge of official interest and research activity. Between 1999 and 2000, the House of Lords Select Committee on Science and Technology undertook an inquiry into CAM (House of Lords Select Committee on Science and Technology 2000b). Of significance to acupuncture was the pragmatic stance taken when considering whether or not patients should be provided access and research undertaken in the absence of plausible explanations for the therapy's mechanisms of action. The opinion of the committee was that the safety of the therapy should determine continued public access and not the strength of its evidence base (House of Lords Select Committee on Science and Technology 2000a). The Select Committee also advised that evidence of a therapy's efficacy determine whether or not wider access via the National Health Service is provided, whatever the controversy over its underlying mechanisms. And, like the European Council (European Commission 1998) and the British Medical Association (British Medical Association 2000), the House of Lords recognised the urgent need for research that answers not only questions of efficacy and safety, but investigates patients' satisfaction with CAM.

1.5.1 Acupuncture and safety

Patients generally perceive acupuncture to be more natural than conventional medicine and have fewer concerns about adverse events. A systematic review of prospective studies of acupuncture safety carried out in other parts of the world, found only two cases of pneumothorax and two cases of broken needles in a quarter of a million treatments (Ernst & White 2001). In the UK, two prospective surveys were undertaken to systematically examine both the rate and nature of adverse events (MacPherson et al. 2001b; White et al. 2001). Similar methods were used to enable comparability. In the first survey involving 576 British Acupuncture Council Members, there were no serious adverse events associated with 34,407 treatments: that is no patients required subsequent hospitalisation (MacPherson et al. 2001b). This outcome corresponds with the other survey involving 31,822 consultations with doctors and physiotherapists who performed acupuncture (White et al. 2001). The

probability of a serious adverse event occurring in 10,000 treatment episodes was found to lie between 0 and 1.1 (95% CI), a rate classified as minimal (British Medical Association Ethics 1993). Both surveys also reported a rate of 14 per 10,000 acupuncture consultations for minor transient but significant adverse events (such as nausea or fainting). A limitation of these studies is the probable existence of reporter bias, the possibility that some patients who experienced an adverse event following a treatment did not return to report the event or complete their course of acupuncture, and that by restricting the survey to immediate complications of treatment, no longer-term deleterious effects on the patient's condition could be identified. The results from a postal survey of prospectively identified acupuncture patients to establish the type and frequency of adverse events they experienced and attributed to their acupuncture treatment, therefore provides useful triangulation (MacPherson et al 2004). This survey involved patients who had received acupuncture treatment from members of the British Acupuncture Council. The adverse event rate reported by patients at follow-up (n=6348/9408) in relation to the 30,196 consultations during the three-month period covered by the survey was 350 per 10,000 consultations (95% CI 330 to 370), and is considerably higher than that reported by practitioners in the surveys reported above. The most common events reported were severe tiredness and exhaustion, pain at the site of needling, and headache. Ten percent of adverse events were due to treatment accidents (moxabustion burns to the skin, needle left in after treatment, electro-acupuncture with too strong a sensation and inserted needle breaking), and is again higher than that reported by practitioners. Serious adverse events that patients associated with acupuncture treatment were rare (0.5 per 1000 patients over three months) and seemingly quite tenuous (falling asleep whilst driving two days after an acupuncture treatment) or indicative of a healing response (an aggravation of existing symptoms followed by an overall improvement). Thus this survey adds additional evidence to support the general finding that acupuncture is safe in competent hands.

1.5.2 Current research of acupuncture for gynaecological and obstetric conditions

Public endorsement, an official call for evidence, evidence of safety, and the subsequent availability of funding has created the social climate necessary for acupuncture research in the UK, a milieu paralleled in other western nations. At the

time of writing, the first randomised clinical trial of acupuncture and Chinese herbs for endometriosis was being planned in Oregon, America (Hammerschlag et al. 2000); a trial of acupuncture for primary dysmenorrhoea was underway in Adelaide, Australia for the first time (Smith 2000); the first UK study to assess whether acupuncture/ transcutaneous electrical nerve stimulation can induce ovulation in women with polycystic ovary syndrome was taking place in Oxford, UK (Kennedy 2002); and the first exploratory trial of acupuncture for the treatment of patients with menorrhagia in primary care in the UK was drawing to a close (Longridge et al. 2001) - this trial was undertaken by the candidate for the purposes of a PhD degree in Health Sciences.

1.6 The ACUMEN programme of research – embracing complexity

The York and Selby Acupuncture for Menorrhagia Exploratory Trial (ACUMEN) with nested qualitative study and complementary GP survey was designed in 2000. At its' core lay the commitment to design a rigorous study that would evaluate acupuncture as practiced by members of the British Acupuncture Council and so allow for individualised points prescriptions that might change over time, the use of co-interventions, development of a beneficial therapeutic relationship, and giving of lifestyle advice. The evaluation of everyday clinical practice was deemed essential to preserve the integrity of the therapy and ensure the findings would be relevant to the acupuncture profession.

The guidelines produced earlier that year for the evaluation of complex interventions to improve health by the Medical Research Council (2000) proved invaluable, validating the pragmatic approach taken and informing the study design. The important message for ACUMEN was that when at an exploratory stage a study can benefit greatly from the use of multiple methods, as these allow the question to be investigated from differing perspectives to give a more comprehensive understanding of the phenomenon. The guidelines highlighted the importance of understanding what personal, professional or organisational factors might make an intervention more or less acceptable. They also described a role for qualitative research in eliciting the nature of the complex intervention delivered. This expansive approach to researching

complex healthcare interventions was considered entirely consistent with the aims of ACUMEN in that it acknowledges and embraces their complexity.

1.7 Structure of the thesis

The thesis contains six chapters, of which this introduction is the first.

In the second chapter, the findings from a descriptive systematic review of the evidence for acupuncture in the treatment of menorrhagia and the related idiopathic conditions primary dysmenorrhoea and irregular menstruation are given. This represents the first attempt to synthesise evidence for the effectiveness of acupuncture in the treatment of menorrhagia.

Chapter Three presents the findings from the York and Selby Acupuncture for Menorrhagia Exploratory Trial (ACUMEN), which is perhaps the first rigorously designed and conducted randomised trial of acupuncture for a menstrual disorder, and certainly for menorrhagia. The feasibility of and justification for a definitive trial to evaluate the relative clinical and cost effectiveness of acupuncture in primary care is considered.

Chapter Four presents the first longitudinal, qualitative study of menorrhagia patients receiving traditional acupuncture within the confines of a trial. From the findings, an attempt is made to understand the factors that led women to choose to accept the offer of acupuncture in ACUMEN, the way in which treatment acceptability changed over time, and how these findings might inform the design and conduct of future trials.

Chapter Five focuses upon the ACUMEN GP postal questionnaire survey. This survey adds to the evidence that exists for factors that hinder and help active patient recruitment by GPs within a trial. It provides evidence of specific relevance to trials of acupuncture for menorrhagia and may inform other trials of acupuncture for conditions *not* characterised by pain and for which there is no credible evidence base.

Chapter Six summarises the findings from all of the above and draws the thesis to its end providing suggestions for possible future research in this area.

Chapter Two

A Descriptive Systematic Review: Acupuncture for the Treatment of Menorrhagia

1. Background and Introduction

1.1 The systematic review

Established in 1994, the NHS Centre for Reviews and Dissemination is a UK based organisation with a primary function of conducting systematic reviews to inform health care policy. The Cochrane Collaboration was founded a year earlier in 1993, and is a respected international research network that specialises in the preparation and dissemination of high quality systematic reviews on the effectiveness of healthcare interventions (Olsen et al. 2001). '*A systematic review is an overview of scientific studies that uses explicit, systematic and therefore reproducible methods to locate, select, appraise and synthesise relevant and reliable evidence*' (Russell et al. 1998). Such reviews are able to minimise the delay between the discovery of new effective treatments and their implementation in practice (Greenhalgh 2001). In addition, a rigorous systematic review provides researchers with comprehensive information by which to identify gaps in the evidence base, and is therefore an invaluable first step prior to embarking upon new primary research. The studies included in systematic reviews are typically limited to well-conducted randomised trials that are selected or rejected according to clear criteria (Clarke & Oxman 2000). However, in cases where it is known that few or no RCTs exist, the assessment of quasi-experimental and observational studies is of value (Goodman 1993; Jadad 2000; White et al. 2002). The purpose of such descriptive systematic reviews is often to stimulate improvements in research and provide a rationale for future research, rather than synthesise outcomes (Light & Pillemer 1984).

1.2 Reviews to determine the effectiveness of acupuncture for gynaecological conditions

A recent Cochrane review to determine the effectiveness of transcutaneous electrical nerve stimulation and acupuncture for primary dysmenorrhea reported on just two

RCTs of acupuncture (Proctor et al. 2002). A briefing paper entitled *Gynaecology and Acupuncture: the evidence for effectiveness* and produced by the Acupuncture Resource Research Centre (ARRC) reported on two studies of menorrhagia, one of amenorrhoea and seven studies of acupuncture for primary dysmenorrhoea (Longridge nee Gamon et al. 2000). These 10 studies included the 2 RCTs selected for the Cochrane review (Helms 1987; Thomas et al 1995). It is worth noting that both these trials were conducted in the West, whilst seven of the remaining eight case-series were conducted in China. Socio-political influences are thought to account for the scarcity of adequately controlled trials from China (Birch & Felt 1999c). For example, until recently, the use of placebo controls was illegal. A key issue, however, is the Chinese acupuncture community's limited understanding of bio-statistical methodology, which has led to outcomes being subjectively assessed by the practitioner-researcher as having, for example, *"markedly improved"* or *"failed to respond"*, and to the use of scientifically meaningless statistics (e.g. counts and averages). The strength of these studies is considered to be the use of authentic and credible acupuncture treatment protocols and large patient groups, whilst inadequate treatments and small sample sizes are a common criticism of the often more rigorously designed Western studies (Birch 1997; Longridge nee Gamon et al. 2000; Hammerschlag & Morris 1997). A descriptive systematic review would provide useful information about the strengths and weaknesses of overall findings, and provide guidance in how best to interpret the findings.

1.3 Publication bias and the grey literature

It is clear from the ARRC's briefing paper that there is a scarcity of available research even when methodological criteria for selection is widened (Longridge nee Gamon et al. 2000). This is partly because very few of the studies published in the hundreds of acupuncture journals in China, Japan or Korea are translated and published in Western acupuncture journals. It is also a reflection of the poor reputation Chinese and Asian work has among Western scientists, because of the unusually positive results cited, a situation that has led to studies published in Asian medical journals in general being largely excluded from mainstream biomedical information systems such as MEDLINE (Birch & Felt 1999c). In a recent systematic review of controlled clinical trials of acupuncture in the Japanese literature, Tsukayama and Yamashita

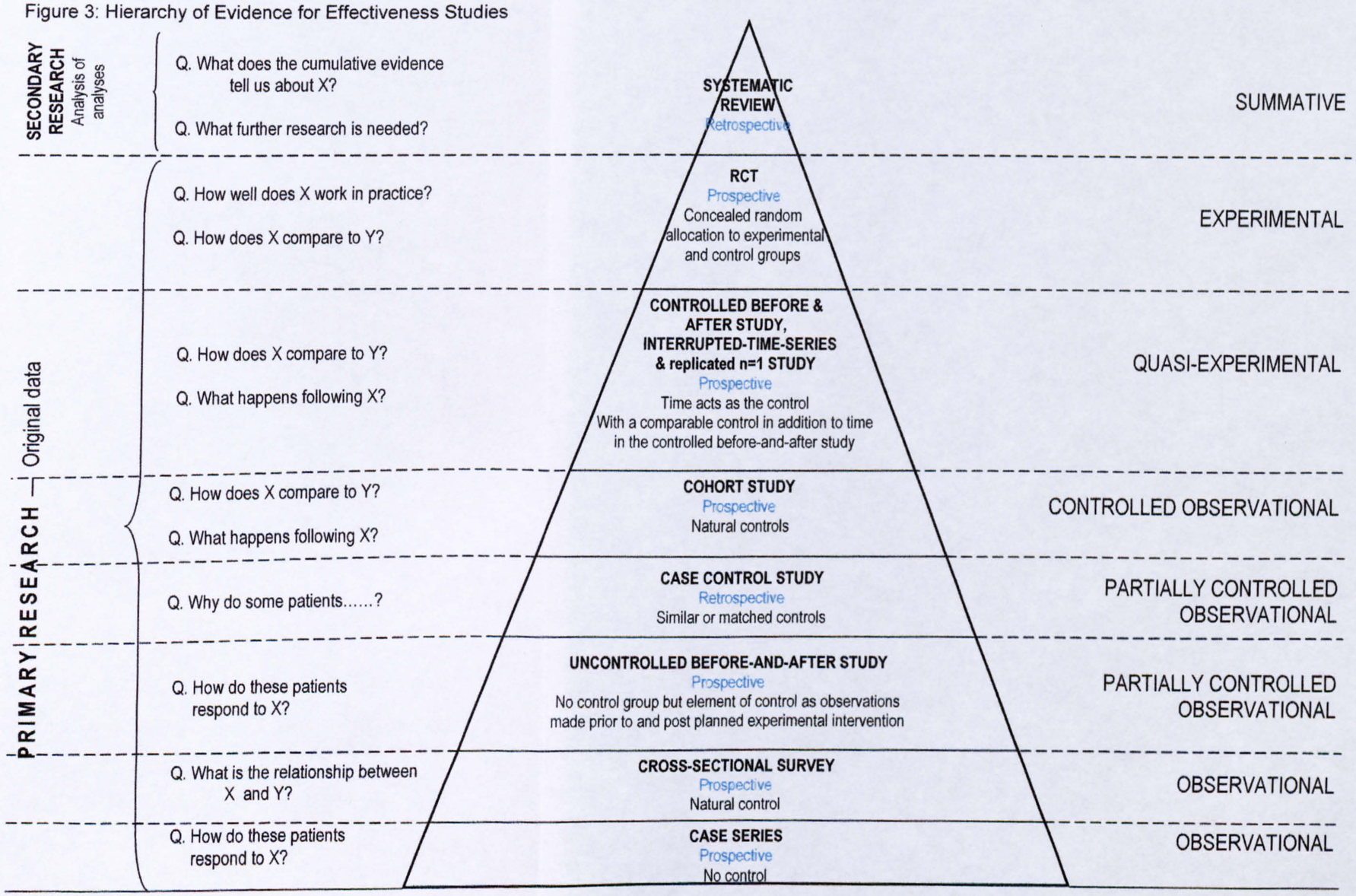
(2002) found that of the 34 trials located only two were listed in MEDLINE or the Cochrane Library. Similarly, the thousands of studies of acupuncture for gynaecological conditions published in the hundreds of Chinese traditional medicine journals currently remain practically hidden (Mrs Mingming Zhang 2003). The founding of a Chinese Cochrane Centre in 1999, with a goal of translating RCTs into English for inclusion in the Cochrane Central Register of Controlled Trials, is a positive first step to making this body of evidence available to more researchers (Zhang 2003). Moreover it is an undertaking that is suited to the Cochrane Collaboration, as the preparation of high quality translations to provide rigorously consistent databases of information demands an approach similar to that required for systematic reviews: Notably the independent translation of studies by more than one translator and the review of these translations by a suitably qualified advisory group. Indeed, such collaborative efforts are essential to promote a more thorough scientific evaluation of acupuncture. It is hoped that there will be similar strategies to facilitate access to the Japanese literature, not least because the quality of this research has improved so dramatically since mid-1990 and the publication of the WHO Guidelines for Clinical Research on Acupuncture (Tsukayama & Yamashita 2002).

1.4 Hierarchy of effectiveness studies

Figure three illustrates a hierarchy of evidence for effectiveness and was drawn from a number of sources (Crombie & Davies 1998; Earl-Slater 2002; Greenhalgh 2001; NHS Centre for Reviews and Dissemination 2001; Thomas & Fitter 2002). This hierarchy reflects the degree to which different study designs are susceptible to bias or deviations from the truth. For example, high-quality randomised controlled trials are widely perceived as the gold standard primary research design for evaluating effectiveness, because they increase the comparability of the groups being compared and so minimise confounding of the intervention of interest by differences in known and unknown prognostic factors between groups. It is for these reasons that systematic reviews tend to concentrate exclusively on randomised trials.

Here controlled-before-and-after studies, interrupted-time-series and replicated n=1 studies are defined as quasi-experimental, as they entail many of the design features of experimental research – standard validated outcome measures, complete follow-up,

Figure 3: Hierarchy of Evidence for Effectiveness Studies



assessor blinding and intention to treat analyses etc. - bar random allocation to groups. Instead time acts as the control, and trends in the dependent variable are observed over time. A stable trend is established via repeated pre-intervention testing, and outcomes are established only after repeated post-intervention testing. The controlled before-and-after study also involves the use of a comparable control, such as patients at a neighbouring GP practice to the one where the experimental intervention is to be introduced. And ideally, quality assurance is provided to attest to the validity of the findings by, for example, trialists seeking to identify confounding from changes in practice organisation or due to the influence of a new GP.

Similarly, the cohort study observes those features of experimental research it can accommodate (e.g. standard validated outcome measures, complete follow-up, assessor blinding and intention to treat analyses), but is lower on the hierarchy and defined here as controlled observational because it involves the use of naturally occurring controls, perhaps in different geographical locations. Here, not randomising participants seriously compromises the comparability of the groups and so weakens the design and the integrity of the results. The cohort study is also particularly open to selection bias, but by documenting known and possible prognostic variables it is possible to better balance the groups during the analysis and provide some measure of quality assurance (Walach et al. 2002).

Case-control studies are conducted retrospectively and test to see if there is an association between the intervention and the outcome of interest. By using matched controls, it aims to minimise flaws in the design that could bias the measurement of association (Downs & Black 1998), but it is impossible to control for everything, meaning these studies are particularly vulnerable to bias from confounding and are defined here as partially controlled observational.

The uncontrolled before-and-after study involves simply observing the effects of a planned intervention upon a population of patients and collecting data on outcomes. There is some element of control as observations are made before and after the planned experimental intervention, which is why it is also classed as partially controlled observational, but it is a design that is vulnerable to a number of threats to validity, including that from observer bias, confounding, spontaneous recovery, and

the Hawthorne effect (where patients respond positively to study participation), which is why it is lower down on this hierarchy.

The cross-sectional survey involves the collection of data on a section of the target population at a single point in time. It therefore provides only indirect evidence about the effects of an intervention upon a population, and great caution must be taken when drawing conclusions about change, as this study design is particularly vulnerable to confounding. Consequently this design is classed here as observational. Case-series are also defined as observational, as they merely describe the way in which one or more patients experienced an intervention and the temporal changes recorded: studies without a comparison group do not allow conclusions about effectiveness to be drawn. Indeed, the case-series is highly susceptible to bias, especially that arising from the way the cases were obtained. Consequently they are at the bottom of the hierarchy and their findings commonly treated with "caution bordering on suspicion" (Crombie & Davies 1998). Thus, it is not uncommon for textbooks on evidence-based health care and the medical literature to omit the case-series completely (Bowers et al. 2001; Gray 1999; Marchevsky 2000). Yet, they do have a role to play, as they can:

I. Enable physicians to traverse the many grey zones of practice,
II. Provide valuable evidence when considering the design, primary question and outcomes for a clinical trial, and
III. Be relatively inexpensive, simple and expedient to conduct (Earl-Slater 2002; Greenhalgh 2001; Grimes & Schulz 2002).

1.4.1 Evidence included in systematic reviews

It is not uncommon for researchers to need to collate evidence from trials, controlled before-and-after studies, cohort, and case-control studies in reviews of healthcare interventions, and there are recognised methods to do this (Downs & Black 1998). What is less common is the planned inclusion of findings from studies at the very bottom of this research hierarchy, namely the uncontrolled before-and-after study, cross sectional survey and case-series. Yet, in reviews of acupuncture for menorrhagia, if clinicians and researchers are to be provided with a comprehensive

assessment of the evidence that currently exists, the criteria must be further extended to include data from all effectiveness studies.

1.5 Aim of the systematic review

The aim of this systematic review is to identify, sort and rationalise evidence from primary research that demonstrates, or generates hypotheses about, the effectiveness of acupuncture for menorrhagia *[Figure 3]*. To this end, the review will also consider findings from studies of acupuncture in the treatment of the commonly associated idiopathic menstrual disorders primary dysmenorrhea and irregular menstruation. And to further assure the comprehensiveness of the review findings, acupuncture studies of other gynaecological conditions will be included where changes in volume of menstrual loss, cycle regularity, and pain at menstruation are outcomes of interest.

1.6 Research questions to be investigated

This systematic review of acupuncture for menorrhagia asks:

I. *"Is there evidence of effectiveness for acupuncture in the treatment of patients with menorrhagia and the related idiopathic menstrual disorders primary dysmenorrhea and irregular menstruation?"*

II. *"Does the evidence justify conducting more rigorous research to assess the benefits of acupuncture for patients with menorrhagia?"*

2. Methods

2.1 Search strategy for identification of studies

2.1.1 The electronic databases

The search strategy in Table 4 was used for the identification of records that describe (or might describe) studies of acupuncture in the treatment of menorrhagia, dysmenorrhea or irregular menstruation. It was adapted from those developed for the Cochrane Menstrual Disorders and Subfertility Group (Farquhar et al. 2002) *[Appendix 1A]*. This adaptation involved refining the scope of the search by omitting

Table 4: Search strategy 04.02.03 - AMED Version

1. exp Menstruation Disturbances/
2. (menstrual$ adj5 disturbance$).ti,ab.
3. Dysmenorrhea/
4. dysmenorrh$.ti,ab.
5. (painful adj5 menstrua$).ti,ab.
6. (painful adj5 period$).ti,ab.
7. Pelvic Pain/
8. (pelvic adj5 pain).tw.
9. (menstrual$ adj5 bleed$).ti,ab.
10. (heavy adj5 menstrua$).ti,ab.
11. (dysfunctional adj5 uter$).ti,ab.
12. Menorrhagia/ or "Menorrhagia and Metrorrhagia"/
13. menorrrhagi$.ti,ab.
14. Metrorrhagia/
15. metrorrhag$.ti,ab.
16. Premenstrual Syndrome/
17. PMS.ti,ab.
18. Premenstrua$.ti,ab.
19. Pre menstrual$.ti,ab.
20. (iron adj5 anaem$).ti,ab.
21. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20
22. exp FERTILITY AGENTS, FEMALE/
23. Infertility, Female/
24. infertility therapy.tw.
25. exp Reproduction/
26. subfertil$.tw.
27. infertili$.tw.
28. (ovulat$ adj5 induct$).tw.
29. exp Menstrual Cycle/
30. exp AMENORRHEA/
31. anovulat$.tw.
32. ovar$.tw.
33. 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32
34. 21 or 33
35. exp Acupuncture/
36. exp Acupuncture Therapy/
37. acupuncture$.tw.
38. 35 or 36 or 37
39. 34 and 38
40. limit 39 to female
41. limit 40 to human
42. limit 41 to English language
43. 41 not 42

subject headings and keywords most unlikely to yield relevant studies, such as "menopause", "abdominal hysterectomy" and "sperm". Terms to identify acupuncture studies were included, and terms designed to limit the search according to study type, such as "controlled study" and "double blind procedure" were omitted.

It is worth noting that the subject headings and keywords used by the author to search for acupuncture studies correspond with those used for a recent Cochrane review of acupuncture in the treatment of primary dysmenorrhea (Proctor et al. 2002) *[Appendix 1B].* These search terms yield studies that have defined the intervention as acupuncture, acupressure, acupuncture analgesia, ear-acupuncture, electro-acupuncture or moxibustion, and describe acupuncture points or meridians.

The strategy in Table 4 was used to search AMED, MEDLINE, EMBASE, CINAHL and CISCOM electronic databases 4th February 2003 using Ovid software. This made a search of the specialist acupuncture database ARRCBASE superfluous, as ARRCBASE contains only articles relevant to Chinese medicine and acupuncture and listed on AMED or MEDLINE.

AMED

The Allied and Complementary Medicine (AMED) bibliographic database is produced by the Health Care Information Service of the British Library. The indexing policy is to select content from journals relevant to several professions allied to medicine, complementary medicine (CAM) and palliative care. In this way it performs a vital role as records are often sourced from journals not indexed by other biomedical sources, such as MEDLINE and EMBASE. This has meant that it is no longer necessary to hand search specialist acupuncture journals published after 1985 (including the Journal of Chinese Medicine, which has been the foremost English language journal dedicated to professional and student level information on the entire field of Chinese medicine for over 20 years, and the European Journal of Oriental Medicine, which has been the journal of the British Acupuncture Council since 1993).

MEDLINE

MEDLINE is the computer version of Index Medicus. It is compiled by the U.S. National Library of Medicine and is the world's most comprehensive source of life sciences and biomedical bibliographic information, indexing over 7,300 different publications and being updated weekly.

EMBASE

The Excerpta Medica database (EMBASE) produced by Elsevier Science, is a major biomedical and pharmaceutical database indexing over 3,500 international journals, including those in the fields of alternative medicine and clinical and experimental human medicine.

CISCOM

CISCOM is the Research Council for Complementary Medicine's (RCCM) database, and is sourced from the British Library and MEDLINE databases, with additional data provided from citation tracking and the RCCM's researcher network. The search of this specialist database was conducted by RCCM staff on behalf of the author, and according to the above search strategy, for a fee.

CINAHL

The Cumulative Index to Nursing and Allied Health Literature (CINAHL) is a bibliographic database indexing from just over 1,200 journals, including a limited number of journals related to acupuncture (e.g. the Journal of Chinese Medicine, but not the American Journal of Acupuncture or the International Journal of Clinical Acupuncture).

2.1.2 Additional search initiatives

In addition to these search initiatives, the author searched the Cochrane Library database CENTRAL, the National Research Register, contacted the Chinese Cochrane Centre, and searched the citation lists of all potentially relevant papers and reviews.

The Cochrane Library – CENTRAL

The Cochrane Library consists of a regularly updated collection of evidence-based medicine databases, and includes the Cochrane Central Register of Controlled Trials (CENTRAL). CENTRAL is a bibliography of controlled trials identified by contributors to the Cochrane Collaboration and others, as part of an international effort to hand search the world's journals and create an unbiased source of data for systematic reviews. CENTRAL includes reports published in conference proceedings

and in many other sources not currently listed in MEDLINE or other bibliographic databases.

National Research Register (NRR)

The NRR is a register of ongoing and recently completed research projects funded by, or of interest to, the United Kingdom's National Health Service. It also includes entries from the Medical Research Council's Clinical Trials Directory, and details of reviews in progress collected by the NHS Centre for Reviews and Dissemination.

2.2 Selection criteria

The predetermined selection criteria for the review were as follows:

1. Evidence from primary research *[Figure 3]:*
    - Randomised controlled trial
    - Controlled before-and-after study
    - Interrupted time-series
    - N=1
    - Cohort study
    - Case control study
    - Uncontrolled before-and-after study
    - Cross-sectional survey
    - Case-series
2. Primary research that has sought to assess the effectiveness of acupuncture in the treatment of menorrhagia, primary dysmenorrhoea and irregular menstruation,
3. Primary research that has sought to assess the effectiveness of acupuncture in the treatment of gynaecological conditions where changes in volume of menstrual loss, cycle regularity, and pain at menstruation are outcomes of interest.

2.3 Quality assessment

The assessment of quality or validity is the process whereby studies are graded according to the reliability of their results, so that they can be given appropriate weight in the synthesis and when drawing conclusions (NHS CRD1996). The range of study types included in this review necessitated the preparation of a number of

UNIVERSITY OF YORK LIBRARY

suitable checklists for methodological quality that ran in tandem to a checklist to assess the quality of interventions in studies of acupuncture. These are summarised in Table 5 and described in sections 2.3.1a, b and c below.

Table 5: Study type and corresponding checklist

Study Type	Classification	Review checklist	Source(s)
RCT	Experimental	1a. Checklist to assess the methodological quality in trials of acupuncture for menstrual disorders 1b. Checklist to assess the quality of interventions in studies of acupuncture	(1) CONSORT statement (2) CRD 2001 guidelines (3) Downs & Black checklist (4) STRICTA statement
Time series studies o Controlled before-and-after study o Interrupted-time series o N=1	Quasi-experimental	2a. Checklist to assess the methodological quality of time series studies, cohort and case control studies of acupuncture for menstrual disorders 2b. *See 1b*	Downs & Black with additional items to check for: o Stable trends o Comparable control o Quality assurance
Cohort study	Controlled observational	*as above*	
Case control study	Partially controlled observational	*as above*	
Uncontrolled before-and-after study	Partially controlled observational	3a. Checklist to assess the methodological quality in observational studies of acupuncture for menstrual disorders 3b. *See 1b*	(5) EBMWG guidelines CRD 2001 guidelines Downs & Black checklist CONSORT
Cross-sectional survey	Observational	*as above*	
Case-series	Observational	*as above*	

References
(1) Altman DG, Schulz K F, Moher D, Egger M, Davidoff F, Elbourne D, & for the CONSORT Group 2001, "The revised CONSORT statement for reporting randomised trials: explanantion and elaboration", *Ann Intern Med*, vol. 134, no. 8, pp. 663-694.; Moher D, Schulz K F, Andersch B, & for the CONSORT Group 2001b, "The CONSORT Statement: revised recommendations for improving the quality of reports of parallel-group randomised trials", *Lancet*, vol. 43, no. 11, pp. 1191-1199; Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, Pitkin R, Drummond R, Schulz KF, Simel D, & Stroup DF 1996, "Improving the quality of reporting of randomised controlled trials", *JAMA*, vol. 276, no. 8, pp. 637-639 (2) NHS Centre for Reviews and Dissemination, U. o. Y. 2001, *Undertaking systematic reviews of research on effectiveness: CRD's guidance for those carrying out or commissioning reviews.* (3) Downs SH & Black N 1998, "The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions", *Journal of Epidemiology and Community Health*, vol. 52, pp. 377-384 (4) MacPherson H, White A, Cummings A, Jobst K, Rose K, Niemtzow R, & for the STRICTA group 2001, "Standards for reporting in controlled trials of acupuncture: the STRICTA recommendations", *Complementary Therapies in Medicine*, vol. 9, pp. 246-249 (5) Randolph A, Bucher H, Richardson WS, Wells G, Tugwell P, & Guyatt G 1994, "Prognosis," in *Users' guides to the medical literature: a manual of evidence-based clinical practice*, The Evidence Based Medicine Working Group - Guyatt G & Rennie D, eds., pp. 141-154.

When developing or adapting these checklists, the candidate took into consideration how they would each constitute one tool in a range of instruments to be used in a single review. Thus to alleviate respondent burden, efforts were made to use a similar layout and language to that in the Downs and Black (1998) throughout. For instance,

studies were assessed as either fully, partially, or not meeting the quality criteria specified by each item, and in the case of the checklist for observational studies at the bottom of the research hierarchy, the 14 items were organised across the subscales: reporting, external validity, confounding and bias.

Table 6: Known confounding factors specific to studies of menstrual disorders

	*Known confounders or prognostic variables	Information for raters
1	Age	Adolescent girls and women aged over 40 years are at greater risk of anovulatory menorrhagia, which gives the heaviest bleeds. Ovulatory menorrhagia is almost exclusively confined to women in their mid-reproductive years and aged between 20 and 40 years.
2	Contraception (OCP and IUCD)	A copper-containing or inert intra-uterine device is often the principle factor in the development of some women's menorrhagia, whilst the contraceptive pill is often able to reduce menstrual loss and pain.
3	Parity	Parous women tend to have heavier, less painful menses than non-parous women.
4	Obesity	Women who become very overweight are more prone to developing anovulatory menorrhagia, which gives the heaviest bleeds

**Shaw R, Soutter WP, Stanton SL, 2002 "Gynaecology", 3rd Edition, Churchill Livingstone*

The majority of the items in each of the three checklists for methodological quality could have been asked of any relevant randomised or non-randomised study *[Tables 7, 8 & 10]*, whilst two were topic sensitive: one requiring that raters were provided with information about relevant confounders *[Table 7, item 3; Appendix 3, item 5; Table 10, item 7]*; and the other linked to the checklist to assess the validity of interventions in trials of acupuncture *[Table 7, item 4; Table 8, item 4; Table 10, item 3]*. For the purposes of this review, the known confounding factors were those listed in Table 6, in addition to those generally accepted to affect health: socio-economic group, ethnicity, severity and chronicity. And the question that asked whether or not the interventions of interest are clearly described became a summary measure derived from the score the study achieved on the checklist for the quality of interventions in studies of acupuncture. Thus, with an "excellent" score of 75% and above on the intervention checklist, the question was answered "yes", with a "good" score of 50% and above "partially" and with a "poor" score below 50% "no". In this way, the candidate sought to enhance the reliability of rater responses to this item.

Table 7: RCT checklist items, rationale for item inclusion, response category construction and scores

	Checklist items	**Rationale**	**Scores**		
			Yes	**No**	**Partial**
1	Is the aim of the study clearly described?	CONSORT / D&B	2	0	1
2	Are both inclusion and exclusion criteria specified?	CONSORT / CRD	2	0	1
3	Are the distributions of principal confounders in each group of subjects to be compared clearly described? *(A list of principal confounders is provided: eg. age, parity, contraception, obesity, severity and chronicity (duration of symptoms).)*	CONSORT CRD / D&B	2	0	1
4	Are the interventions of interest clearly described and their timing? *(Where a score of 75% or more is achieved on the checklist for the quality of interventions in trials of acupuncture, this question should be answered yes. Answer partially for a score of 50% or above, and no for a score below 50%.)*	D&B / CONSORT / STRICTA	2	0	1
5	Are primary and secondary outcome measures clearly defined? *(This should include a statement about minimum important differences).*	CONSORT / CRD	2	0	1
6	Is justification for the sample size given? *(If the effect size is given and a power calculation was undertaken prior to the study answer yes. Answer partial if only effect size was discussed* ***or*** *only a power calculation was undertaken).*	D&B	2	0	1
7	Is the method of randomisation secure in principle? *(Answer yes if is an on-site, coded computer system that gives allocations only after inputting an enrolled participants details; is a remote-telephone system; or envelopes sequentially numbered, sealed and opaque. Answer no if randomisation involves a "list" or "table" to allocate assignments; or for "envelopes" or "sealed envelopes" without giving further details.)*	CONSORT / CRD	2	0	1
8	Are the methods for allocation concealment adequate? *(Answer yes if the paper convinces you that allocation cannot be predicted. Answer partially if is partially convincing).*	CONSORT / CRD	2	0	1
9	Are those assessing the outcomes blinded to group assignment?	CRD / D&B	2	0	1
10	Is baseline comparability assessed?	CONSORT / CRD	2	0	1
11	Is compliance assessed?	D&B	2	0	1
12	Is the attrition rate acceptable? *(Answer yes if less than 20%, partial if between 20 and 30% and no if more than 30%)*	D&B	2	0	1
13	Has intention to treat analysis been carried out? *(Answer yes if all participants were analysed according to their original group assignment regardless of what subsequently occurred.)*	CONSORT / CRD	2	0	1
14	Are the statistical analyses appropriate? *(Answer yes if yield an estimate of effect on primary and secondary outcome measures; include a point estimate and measure of precision; give actual p values; and give absolute numbers as well as percentages)*	CONSORT / CRD	2	0	1
15	Are there enough summary data and descriptive and inferential statistics to permit alternative analyses and replication?	CONSORT / D&B	2	0	1
16	Do the findings support the conclusions? *(Do the authors note sources of bias and imprecision and discuss the issue of external validity? Do they give guidance on interpretation of the evidence in light of the totality of evidence available?)*	CONSORT	2	0	1
			Maximum quality assessment score of 32		

CONSORT - *Begg C et al. 1996 "Improving the quality of reporting of randomised controlled trials", JAMA, vol. 27, no.8, pp. 637 –9.* **CRD**- *NHS CRD, University or York, 2001 "Undertaking systematic reviews of research on effectiveness: CRD's guidance for those carrying out or commissioning reviews"* **D&B** - *Downs SH, Black N, 1998 "The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health interventions", Journal of Epidemiology and Community Health, Vol.52, pp.* 377-384 **STRICTA** - *MacPherson H et al. 2001 "Standards for reporting interventions in controlled trials of acupuncture: the STRICTA recommendations"* Complementary Therapies in Medicine, vol. 9, 246-9

2.3.1a Checklist to assess the methodological quality in trials of acupuncture for menstrual disorders

The checklist used for the assessment of trials had 16 items and was developed with reference first, to the original and recently revised Consolidated Standards for Reporting Trials (CONSORT) statement (Altman et al. 2001; Begg et al. 1996; Moher et al. 2001b); second, the 2001 CRD guidelines (NHS Centre for Reviews and Dissemination 2001); and third, the Downs and Black (1998) checklist for randomised and non-randomised studies.

The CONSORT statement is a consensus document that details the information required for judgements to be made about the external and internal validity of a trial, and is readily adapted into a succinct checklist for reviews (Jadad et al. 1996; Moher et al. 1995; Moher et al. 2001a), whilst the CRD guidelines outline key questions that should be asked when assessing trial validity. Undoubtedly these documents were the most influential when developing this checklist for trials, but certain items, or the wording for certain items, were included from the Downs and Black (1998) checklist where they gave greater clarity of meaning. The Downs and Black (1998) checklist was used in this review for assessing time-series, cohort and case-control studies. The decision not to use it for trials and instead to employ a separate trials checklist, was made in the interests of the greater ease of use and clarity of interpretation that a specific and therefore more succinct line of question would afford; the Downs and Black checklist contains almost 30 items, as opposed to the 16 in this checklist for trials.

The items and the relevant reference justifying their inclusion are given in Table 7. The items were arranged under the headings "reporting", "methods" and "results" in the checklist for reviewers *[Appendix 2],* and to better enable the checklist to distinguish between trials, a three-point-scale was used (yes=2, partial=1, no=0) (Streiner & Norman 1995).

Table 8: Changes made to the Downs and Black checklist (1998) for this reviews assessment of time-series, cohort and case-control studies - modifications highlighted in red and auxiliary items in blue

Item	Rationale	Auxiliary and modified checklist items	Scores			
			Y	P	N	NA
3	D&B / IR	Are the characteristics of the patients included in the study clearly described? *(In ~~trials~~ time-series and cohort studies, inclusion and/or exclusion criteria should be given. In case-control studies, a case-definition and the source for controls should be given.)*	2	1	0	
4	D&B / STRICTA	Are the interventions of interest clearly described? *(Treatments and placebo (where relevant) that are to be compared should be clearly described. Where a score of 75% or more is achieved on the Checklist for Acupuncture Quality, this question should be answered yes. Answer partially for a score of 50% or above.)*	2	1	0	
11	D&B / IR	Were the subjects asked to participate in the study representative of the entire population from which they were recruited? *(The study must identify the source population for patients and describe how the patients were selected. Patients would be representative if they comprised the entire source of population, or were an unselected sample of consecutive patients, or a random sample. ~~Random sampling is only feasible where a list of all members of the relevant population exists.~~ Where a study does not report the proportion of the source population from which the patients are derived, the question should be answered as unable to determine.)*	2	1	0	
			Y	N	?	NA
X	IR	In time series studies (i.e. controlled-before-and-after studies, interrupted-time-series and large n=1 studies), was there evidence of stable trends? *(Answer NA if the study was a cohort or case control study)*	1 2	0 0	0 1	-2
17	D&B / IR	In ~~trials~~ time-series and cohort studies, do the analyses adjust for different lengths of follow-up of patients, or in case-control studies, is the time period between the intervention and outcome the same for cases and controls? *(Where follow-up was the same for all study patients the answer should be yes. If different lengths of follow-up were adjusted for by, for example, survival analysis the answer should be yes. Studies where differences in follow-up are ignored should be answered no.)*	1 2	0 1	0 0	
22	D&B / IR	Were the patients in different intervention groups (trials and cohort studies) or were the cases and controls (case-control studies) recruited from the same population? *(For example, patients for all comparison groups should be selected from the same hospital. The question should be answered unable to determine for cohort and case-control studies where there is no information concerning the source of patients included in the study.)*	1 2	0 1	0 0	-2
23	D&B / IR	Were the study subjects in different intervention groups (trials and cohort studies) or were the cases and controls (case-control studies) recruited over the same period of time? *(For a study which does not specify the time period over which patients were recruited, the question should be answered as unable to determine.) NA?*	1 2	0 1	0 0	
24	D&B	~~Were study subjects randomised to intervention groups? (Studies, which state that subjects were randomised, should be answered yes, except where method of randomisation would not ensure random allocation. For example, alternate allocation would score no because it is predictable.)~~	~~1~~	~~0~~	~~0~~	
25	D&B	~~Was the randomised intervention assignment concealed from both patients and health care staff until recruitment was complete and irrevocable? (All non-randomised studies should be answered no. If assignment was concealed from patients but not from staff, it should be answered no.)~~	~~1~~	~~0~~	~~0~~	
26	D&B / IR	Was there adequate adjustment for confounding in the analyses from which the main findings were drawn? *(This question should be answered no ~~for trials if: the main conclusions of the study were based on analyses of treatment rather than intention to treat; the distribution of known confounders in the different treatment groups was not described; or the distribution of known confounders differed between the treatment groups but was not taken into account in the analyses. In non-randomised studies~~ if the effect of the main confounders was not investigated or confounding was demonstrated but no adjustment was made in the final analyses ~~the question should be answered as no.~~)*	1 2	0 1	0 0	
Y	IR	Was quality assurance demonstrated? *(To answer yes, the study should state how it had considered, and sought to identify or address, possible bias and confounding from, for example, changes in practice organisation, the influence of a new GP, or the existence of both known and unknown prognostic variables.)*	2	1	0	

References: D&B - Downs SH, Black N, 1998 "The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health interventions", Journal of Epidemiology and Community Health 52:377-384 **STRICTA** - MacPherson H et al. 2001 "Standards for reporting interventions in controlled trials of acupuncture: the STRICTA recommendations" Comp Ther Med 9:246-9 **IR** – Russell I 1999 Methods of Clinical Evaluation module, Department of Health Sciences and Clinical Evaluation, University of York

**Y=yes, P=partially, N=no, ?=unable to determine, NA=not applicable*

2.3.1b Checklist to assess the methodological quality in time-series, cohort and case-control studies of acupuncture for menstrual disorders

The review checklist to assess the methodological quality of time-series, cohort and case-control studies was devised by adapting the Downs and Black (1998) checklist, so that it might assess the quality of quasi-experimental time-series studies, in addition to cohort and case-control studies. The Downs and Black checklist (1998) for randomised and non-randomised studies, is a widely used instrument that provides a measure of external as well as internal validity. Rigorous testing established the reliability of the "reporting" and "internal validity" subscales for this checklist, whilst the "external validity" subscale has received a type of informal validation through the instruments widespread use, often by well respected health scientists (Reeves et al. 2001) *[Appendix 3a*].

For the purposes of this reviews assessment of time-series, cohort and case-control studies, items relevant to randomised trials alone were excluded *[Table 8, items 24 & 25]*, and auxiliary items were included *[Table 8, items X & Y]*, or certain items changed or elaborated upon *[Table 8, items 3, 4, 11, 17, 22, 23, 26]*, to check for stable trends, the use of comparable controls, and evidence of quality assurance. To better enable the checklist to distinguish between studies, scores for the response categories "yes", "partially" and "no" for questions within the reporting section, were scored 2, 1 and 0 respectively, as opposed to scoring only a "yes" response (Streiner & Norman 1995). And the "not applicable" parameter was included to ensure that studies were not tested against inappropriate quality criteria *[Table 8, items X and 22]:* This simply involved generating a composite score and changing the denominator to account for the number of items answered "NA". These modifications are detailed in Table 8, whilst the reviewer's checklist can be found in the Appendices *[Appendix* 3b*]*.

2.3.1c Checklist to assess the methodological quality of uncontrolled before-and-after studies, cross-sectional surveys and case-series studies of acupuncture for menstrual disorders

The review checklist for those studies at the bottom of the research hierarchy *[fig.2]* was developed first, with attention to the guide designed by the Evidence Based Medicine Working Group (EBMWG) to assess articles on prognosis (Randolph et al.

1994) as advised in the CRD guidelines 1996 (NHS CRD 1996) *[Table 9];* second, with reference to the quality criteria for the assessment of observational studies listed in the CRD guidelines (NHS CRD 2001); third with reference to the checklist developed by Downs and Black (1998); and fourth, with attention to the original and recently revised CONSORT statement (Altman et al. for the CONSORT Group 2001; Begg et al. 1996; Moher et al. for the CONSORT Group 2001b) *[Appendix 4]*.

Table 9: Users' guides to an article about prognosis (Randolph et al. 1994)

Are the results valid?
• Was the sample of patients representative
• Were the patients sufficiently homogenous with respect to prognostic risk?
• Was follow-up sufficiently complete?
• Were objective and unbiased outcome criteria used?
What are the results?
• How likely are the outcomes over time?
• How precise are the estimates of likelihood?
How can I apply the results to patient care?
• Were the study patients and their management similar to those in my practice?
• Was the follow-up sufficiently long?
• Can I use the results in the management of patients in my practice?

By basing the items in the checklist upon what the CRD and EBMWG, in particular, deemed to be relevant, important and discriminating, the candidate sought to develop a checklist capable of assessing the desired qualities and covering the relevant domains, and in so doing satisfy the requirements for face and content validity. The 14 items and references justifying their inclusion are given in Table 10. The reporting subscale *[items 1 to 4]* was modelled upon that in the Downs and Black checklist, and assessed whether or not the information in the publication was sufficient to allow a reader to make a fair and accurate assessment of the study findings. In particular, item 2 addressed the tendency for investigators to overstep their data, and assumes that a clear acknowledgement that a descriptive study is *"concerned with and designed only to describe the existing distribution of variables, without regard to causal or other hypotheses"* (Last 1988), is likely to be a reflection of methodological quality. Five of the remaining 10 items addressed the way patients came to be involved in the study (external validity) and the presence of known prognostic factors amongst them (confounding) *[items 5 to 9]*, and five items comprised the bias subscale *[items 10 to 14]* - the ways in which outcomes were measured and analysed. By using a number of items rather than one general item to assess biases arising from either the way patients were selected or outcomes were measured, the candidate

Table 10: Uncontrolled before-and-after studies, cross-sectional surveys & case-series checklist items, rationale for item inclusion, response category construction and scores

No.	**Rationale**	**Checklist items** Scores	**Y**	**P**	**N**	**NA**
1	*CONSORT / D&B*	Is the hypothesis/aim/objective of the study clearly described?				
2	*EBMWG*	In the introduction or discussion, do the authors acknowledge that the study is concerned with and designed only to describe the existing distribution of variables, without regard to causal hypotheses? *(Answer no if authors overstep the data and present results as clear evidence of effectiveness)*	2	1	0	
3	*D&B / CONSORT / STRICTA*	Are the interventions of interest clearly described? *(Treatments should be clearly described. Where a score of 75% or more is achieved on the Checklist for Acupuncture Quality, this question should be answered yes. Answer partial for a score above and no for a score below 50%)*	2	1	0	
4	*D&B*	Are the main findings of the study clearly described? *(Simple outcome data (including denominators and numerators) should be reported for all major findings so that the reader can check the major analyses and conclusions)*	2	1	0	
5	CRD / D&B / EBMWG	Is the process used to select the sample completely defined? *(To answer yes, the study must identify the source population and describe how the patients were selected or recruited)*	2	1	0	
6	CRD / D&B	Are the characteristics of the patients included in the study clearly described? *(To answer yes, the study must give explicit inclusion criteria)*	2	1	0	
7	EBMWG / D&B	Were the patients sufficiently homogenous with respect to prognostic risk? *(To answer yes, the study must describe the distribution of principle confounders such as age, contraception and parity)*	2	1	0	
8	CRD / D&B / EBMWG	Were patients recruited from the same time-period, so avoiding non-contemporaneous bias through changes in definitions, diagnoses, exposures, disorders and interventions over time? *(Where the time-period over which patients were recruited is not stated the question should be answered no)*	2	1	0	
9	CRD	If sub-group comparisons are made, are the distributions of principle confounders in each group of subjects to be compared clearly described? *(E.g. age, parity, contraception, obesity, severity, chronicity and co-morbidity)*	2	1	0	-2
10	D&B / CRD /EBMWG	Are the main outcome measures accurate (valid and reliable)? *(For studies which refer to other work or that demonstrate the outcome measures are accurate, the question should be answered yes. For studies where the outcome measures are clearly described, the question should be answered partially.)*	2	1	0	
11	CRD / D&B	Was an attempt made to blind those measuring the main outcomes of the intervention?	2	1	0	
12	CRD / EBMWG	Was follow-up long enough for important events to occur?	2	1	0	
13	D&B / EBMWG	Was follow-up sufficiently complete? *(This should be answered yes where validity of study is not threatened by a loss that is large in relation to the proportion of patients who did not respond, whose symptoms became worse, or who suffered unacceptable side effects from the intervention. Answer no where losses to follow-up are large or where a study does not report the number of patients lost to follow-up.)*	2	1	0	
14	D&B / CONSORT	Were the statistical tests used to assess the main outcomes appropriate? *(The statistical techniques used must be appropriate to the data. For example non-parametric methods should be used for small sample sizes. Where little statistical analysis has been undertaken but where there is no evidence of bias, the question should be answered yes. If the distribution of the data (normal or not) is not described it must be assumed that the estimates used were appropriate and the question should be answered yes.)*	2	1	0	
Maximum quality assessment score of 28 minus NA scores						

EBMWG – *Randolf A et al. 1994 "Prognosis" In "User's guides to the medical literature: a manual of evidence based clinical practice" Ed. The Evidence Based Working Group* **CRD**- *NHS CRD, University or York, 2001 "Undertaking systematic reviews of research on effectiveness: CRD's guidance for those carrying out or commissioning reviews"* **D&B** - *Downs SH, Black N, 1998 "The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health interventions", Journal of Epidemiology and Community Health 52:377-384* **CONSORT** - *Begg C et al. 1996 "Improving the quality of reporting of randomised controlled trials"JAMA 27; 8:637–9. Altman DG et al. 2001 " The revised CONSORT statement for reporting randomised trials: explanation and elaboration" Ann Intern Med 134;8:663-694. Moher D et al. 2001 "The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials" Lancet 43;11:1191-9* **STRICTA** - *MacPherson H et al. 2001 "Standards for reporting interventions in controlled trials of acupuncture: the STRICTA recommendations" Comp Ther Med 9:246-9*

aimed to increase the sensitivity of the instrument and better enable it to discriminate between studies in these key areas (Streiner & Norman 1995). A three-point-scale was also used (yes=2, partial=1, no=0) to help distinguish between trials. And for trials that did not conduct sub-group analyses, and for whom item 9 was "not applicable", the checklist denominator was changed to avoid generating a misleading or unrepresentative percentage score *[Table 10]*. The candidate and her supervisor piloted this checklist independently on four of the review studies and agreed upon revisions prior to the assessment proper.

2.3.1d Checklist to assess the quality of the interventions in randomised and non-randomised studies of acupuncture

The checklists available, and guidelines for developing checklists, are concerned primarily with how the methodological quality of the trial design can bias outcome (Clarke & Oxman 2000; NHS Centre for Reviews and Dissemination 2001). However, as inadequate acupuncture treatment can also jeopardise trial validity and bias outcome, it is essential that systematic reviews have a method for assessing whether treatments meet basic criteria (Ezzo et al. 2001). Various authors have suggested strategies by which to assess the quality of the acupuncture intervention and give an indication as to the generalisability of the results, but their usefulness has been severely compromised by being biased towards a particular style or school of acupuncture (Birch 1997; Birch 1998; Hammerschlag & Morris 1997; NHS Centre for Reviews and Dissemination 2001).

This review was able to avoid these pitfalls by virtue of the STRICTA recommendations published in 2001 by an international group of acupuncturists from different schools of thought (MacPherson et al. 2001c). Their aim, in line with the CONSORT statement, was to improve the standards for reporting interventions in controlled trials of acupuncture practiced according to a wide range of styles (e.g. western trigger point or traditional acupuncture) and using an equally wide range of methods (e.g. electro-acupuncture and auricular acupuncture) and in the context of both explanatory and pragmatic study designs. By converting the items from these guidelines into a series of questions, the candidate aimed to develop an assessment instrument that would accurately reflect the quality of interventions in acupuncture

studies and satisfy the requirements for face and content validity *[Appendix 5]*. Moreover, because the STRICTA recommendations are congruent to a wide range of acupuncture styles, this checklist also aimed to meet the requirements for model validity or congruity, whereby the scores reflect the quality of the interventions according to the criteria of the particular acupuncture system under investigation (Walach et al. 2002).

This review checklist had a total of 23 items or questions. Each was scored on a three point scale (yes=2, partial=1, no=0) or as Not Applicable (NA). The NA parameter was included to enable studies that either had no control, or did not use co-interventions, to be accurately and fairly assessed *[items 14 & 15, and items 19 to 23]*. This simply involved generating a composite score and changing the denominator to account for the number of items answered "not applicable". The candidate and her supervisor piloted the checklist independently on four review studies and agreed upon revisions prior to the assessment proper.

2.4 Data extraction

The aim of data extraction is to extract necessary data from the selected studies, accurately and without introducing bias (Clarke & Oxman 2000; NHS CRD 2001; NHS CRD 1996). This is best done with a data extraction form, which performs 3 main functions:

- Bridges the gap between the included studies and the systematic review,
- Forms an historical record of the search process,
- Acts as the repository for the data from which the analysis will be made,
- Provides the information needed to construct Tables summarising the studies included in the review.

The data extraction form developed for this review was designed to gather both general information about the publication, and specific information about the study that was directly linked to the review question, the quality assessment checklists and the planned descriptive analysis (NHS CRD 2001) *[Appendix 6]*. For instance, the summary details verifying the study's eligibility for the review asked for both the "gynaecological or obstetric condition treated" and the "outcomes of interest", as a

study of acupuncture in the treatment of infertility that monitored changes in menstrual loss or cycle regularity overtime would be included in this review. The section on study design and methods listed all those quality criteria of relevance when assessing effectiveness studies across the research hierarchy *[Figure 3]*, and the information to be extracted about the study population paid particular attention to subgroups and the distribution of prognostic factors, as it is not unusual for researchers to assess first, overall responses to acupuncture treatment, and second, whether responses differed according to traditional syndrome patterns. The interventions section is directly linked to the corresponding checklist *[3.3.1d]*, and because of the heterogeneous nature of the data to be collected, whilst the form asked for detailed information about the outcome measures, detailed statistical data was not gathered; rather whether or not analyses were appropriate was recorded, along with a summary of results and the rater's view as to whether or not findings supported conclusions.

Assessment

The candidate undertook a full data extraction and she and her supervisor (IR) independently assessed the selected studies using the checklists developed. Where total scores differed by less than three points, the average of the scores was taken; with poor item agreement discussed and the probable cause recorded. Where total scores differed by more than three points, item scores were carefully compared to discern response patterns and identify items producing major discrepancies. Amendments were made to checklist items and guidelines to improve the criteria for assessments and thus the reliability of rater scores. Two independent assessments were considered appropriate for the modest aims and objectives of this descriptive review and exceeds what is regarded as the minimum acceptable practice in systematic reviews, ie a second assessor checking data extraction and quality assessment (NHS CRD 2001).

2.5 Data synthesis

This phase of the review involved the synthesis of extracted data through the tabulation of study characteristics and results to summarise their findings. This approach allowed a descriptive or qualitative assessment of the evidence (NHS Centre for Reviews and Dissemination 2001).

3. Results

3.1 Sources of the identified studies

3.1.1a Results of electronic database searches

In a first step towards the selection of studies for the review, 578 of the 622 references on the literature lists were excluded for reasons one to five below:

1. Duplicate reference,
2. Paper was related to treatment and not treatment evaluation,
3. Paper was a commentary,
4. Study did not involve insertion of acupuncture needles,
5. Paper was related to conditions other than those specified for the review,

Of the 44 papers retrieved, a further 20 were excluded for the reasons listed above, leaving a total of 22 potentially relevant papers, including 1 review that was excluded once its citation list had been searched *[Table 11]*.

Table 11: Search results

Electronic database	**Scope of search**	**References**	**Potentially relevant studies identified from literature lists**	**Potentially relevant studies post-retrieval & excluding duplicates**
AMED	1985 to January 2003	145	26	17
MEDLINE	1966 to January week 2 2003	216	17	2
EMBASE	1980 to 2003 week 5	93	8	2
CISCOM	1966 to January week 5 2003	50	19	1
CINAHL	1982 to December week 4 2002	118	1	0
	TOTAL	**622**	**71**	**22**
	Excluding reviews (n=1)			**21**
	***Including references from additional search initiatives* (n=1)**			**22**

3.1.1b Results from additional search initiatives:

The Cochrane library database CENTRAL dated 2002, Issue 4 was searched for any trials with acupuncture in the title, abstract or keyword, but no additional trials were found. A search of the National Research Register using the keyword acupuncture did not yield any additional studies.

The author contacted the Chinese Cochrane Centre by e-mail to enquire whether it would be possible to conduct a search of Chinese databases to ascertain the number of trials with menorrhagia, metrorrhagia or dysfunctional uterine bleeding in the title. She also asked for information on relevant trials that had been translated into English. Although unable to help for lack of resources, the centre was keen to stress the volume of work that would be available to researchers in the future (Zhang 2003).

A search of the citation lists of all potentially relevant papers, including the one review, supplemented the yield from the electronic databases by one paper. The omission of reference lists in papers translated from Chinese (n=14), or reference lists that had not been translated from Chinese (n=2), may account in part for the limited success of this search initiative.

3.1.1c Included and excluded studies

Eighteen of the 22 potentially relevant papers retrieved were accepted for inclusion in the review (Ge, Ge & Tang 2000; Helms 1987; Jiao 1997; Liang 1990; Liu , Liu & Liu 1992; Ma, Ji, Sha & Liu 1999; Mo , Li , Pu , Xi, Le & Fu 1993; Steinberger & Slagoski 1980; Thomas, Lundeberg, Bjork & Lundstrom-Lindsbedt 1995; Tureanu & Tureanu 1994; Wang, Wang, Jing, Wang 1996; Wang 1987; Wangcheng, Jin, Yimin, Tingfu 1988; Yong-hai & Xiu-li 1995; Yu, Pan, & Wang 1998; Yuqin 1984; Zhan 1990; Zhang & Wang 1994). The table of included studies annexed to the chapter summarises the characteristics of the included studies. The table of excluded studies that follows it details the reasons for the four exclusions (Griffiths 2000; Sha & Ma 2000; Hu 1999; Ji, Ma, Wang & Liu 1999).

3.2 *The included studies*

3.2.1 Distribution of the review studies across the evidence hierarchy

Table 12 shows the distribution of the review studies across the evidence hierarchy and the menstrual condition treated. In summary there are three controlled trials and 15 case-series. Two of the three trials focus upon the treatment of dysmenorrhoea and one the treatment of amenorrhoea. The majority of the studies (12/18) are

concerned with the treatment of dysmenorrhoea, four the treatment of irregular cycles and amenorrhoea and two the treatment of excessive or prolonged menstrual loss.

Table 12: Distribution of the review studies across the evidence hierarchy

	Number in review	Number per condition	Menstrual condition	Study
Controlled trials	**3**	2 1	Primary dysmenorrhoea Post-medication amenorrhoea	Thomas 1995, Helms 1987 Ma 1999
Controlled before-and-after studies, interrupted time-series, n=1	**0**			
Cohort studies	**0**			
Case-control studies	**0**			
Uncontrolled before-and-after studies	**0**			
Cross-sectional surveys	**0**			
Case-series	**15**	2 1 10 1 1	Menorrhagia (DUB and FUB*) Ovulatory dysfunction (amenorrhoea and FUB) *Primary dysmenorrhoea* Post-OCP menstrual *irregularities (long or irregular* cycles and amenorrhoea) Amenorrhoea	Zhang 1994, Wangcheng 1988 Mo 1993 Ge 2000, Yu 1998, Jiao 1997, Wang 1996, Yong-hai 1995, Liang 1990, Zhan 1990, Wang 1987, Yuqin 1984, Steinberger 1981 Tureanu 1994 Liu 1992
Totals	**18**	**18**		

**DUB=dysfunctional uterine bleeding, FUB=functional uterine bleeding – both refer to excessive or prolonged menstrual loss in the absence of organic pathology*

3.2.2 Quality assessment scores

This review did not require the checklist to assess the methodological quality of time-series, cohort and case-control studies that was devised by adapting the Downs and Black (1998) checklist. Rather it involved the checklists devised to assess the methodological quality in (1) trials and (2) case-series studies, and (3) the checklist devised to assess the quality of the interventions in acupuncture studies *[3.3]*.

Methods in trials

The checklist for trials was not piloted before the assessment proper, but performed well with no more than one point between raters' scores *[Table 13]*. The studies by Helms (1987) and Thomas (1995) gained mediocre mean scores of 51 and 55% respectively, whilst the study by Ma (1999) scored poorly (23%). Notably, not one of the studies earned points for a sample size calculation or methods of assignment.

Table 13: Scores for methods in controlled trials of acupuncture for menstrual conditions – article numbers 1 (Helms 1987); 2 (Ma 1999); and 4 (Thomas (1995)

Item	Heading	Question	1	2	4
1	**Reporting**	Is the aim of the study clearly described?	**2** 2	**0** 0	**2** 2
2	**Methods: Protocol**	Are inclusion & exclusion criteria specified?	**1** 2	**1** 1	**1** 2
3		Are the distributions of principal confounders in each group of subjects to be compared clearly described? *(A list of principal confounders is provided: eg age, parity, contraception, obesity, severity and chronicity)*	**2** 2	**1** 1	**2** 2
4		Are the interventions of interest clearly described and their timing? *(Precise details of treatments and placebo (where relevant) that are to be compared and how and when they were actually administered. Where a score of 75% or more is achieved on the Checklist for Acupuncture Quality, this question should be answered yes. Answer partially for a score of 50% or above.)*	**1** 1	**0** 0	**2** 2
5		Are primary & secondary outcome measures clearly defined? *(This should include a statement about minimum important differences).*	**1.5** 1	**1** 1	**1** 1
6		Is justification for the sample size given? *(If the effect size is given and a power calculation was undertaken prior to the study answer yes. Answer partial if only effect size was discussed* **or** *only a power calculation was undertaken).*	**0** 0	**0** 0	**0** 0
7	**Methods: Assignment**	Is method of randomisation secure in principle? *(Answer yes if is an on-site, coded computer system that gives allocations only after inputting an enrolled participants details; is a remote-telephone system; or envelopes sequentially numbered, sealed and opaque. Answer no if randomisation involves a "list" or "table" to allocate assignments; or for "envelopes" or "sealed envelopes" without giving further details.)*	**0** 0	**0** 0	**0** 0
8		Are the methods for allocation concealment adequate? *(Answer yes if the paper convinces you that allocation cannot be predicted. Answer partially if is partially convincing).*	**0** 0	**0** 0	**0** 0
9	**Results: Analysis**	Are those assessing the outcomes blinded to group assignment?	**2** 1	**0** 0	**1** 0
10	I.	Is baseline comparability assessed?	**1** 2	**0** 0	**1** 1
11		Is compliance assessed?	**0** 0	**0** 0	**2** 2
12		Is the attrition rate acceptable? *(Answer yes if less than 20%, partial if between 20 and 30% and no if more than 30%)*	**2** 2	**2** 2	**2** 2
13		Has intention to treat analysis been carried out? *(Answer yes if all participants were analysed according to their original group assignment regardless of what subsequently occurred.)*	**1** 0	**2** 0	**1** 0
14		Are the statistical analyses appropriate? *(Answer yes if: yield an estimate of effect on primary and secondary outcome measures; include a point estimate and measure of precision; give actual p values; & give absolute numbers as well as percentages (10/20, not 50%).)*	**1** 1	**1** 1	**1** 1
15		Are there enough summary data and descriptive and inferential statistics to permit alternative analyses and replication?	**1** 1	**0** 1	**1** 2
16		Do the findings support the conclusions? *(Do authors note sources of bias and imprecision and discuss the issue of external validity? Do they give guidance on interpretation of the evidence in light of the totality of evidence available?)*	**1** 1	**0** 0	**1** 0
	Total	**Maximum score = 32**	**16.5** 16	**8** 7	**18** 17
		Percentage score calculated using average score	**51%**	**23%**	**55%**

Black ink indicates second rater's scores (IR) and blue ink the candidate's
Key: 2=yes; 1=partially; 0=no

Table 14: Scores for interventions in trials of acupuncture for menstrual conditions – article numbers 1 (Helms 1987); 2 (Ma 1999); and 4 (Thomas (1995)

Item	Description	1	2	4
	Acupuncture Rationale			
1	Is the style of acupuncture used defined? *The approach taken should be stated, e.g. traditional Chinese medicine or western medical.*	2 2	2 2	2 1
2	Is the rationale for treatment and whether it is individualised or standardised given? *The rationale for prescribing treatment, for example according to syndrome patterns, segmental levels or trigger points should be stated.*	2 2	0 1	2 1
3	Are literature sources to justify the rationale provided?	2 2	0 0	2 2
	Needling details			
4	Are points used described with standard nomenclature (e.g. ST36, Zusanli) or in terms of anatomical location?	2 2	2 2	2 2
5	Are specific point locations, and whether **unilateral or bilateral**, given for a formula of points? Or, where treatment is individualised, are typical points and the range of points used given?	2 2	1 1	2 2
6	Is the number of needle insertions specified? *A simple total of needle insertions for a formula of points, or as a mean and range where the number of needles varies between patients*	2 2	1 1	2 2
7	Are depths of insertion given (e.g. *cun* or tissue level) and whether standardised or individualised?	0 0	0 0	2 2
8	Are responses elicited described (e.g. deqi or twitch response)?	0 0	0 1	2 2
9	Is needle stimulation described (e.g. manual or electrical)?	1 1	2 2	2 2
10	Is needle retention time specified? *The standard time needles were retained, or the mean and range for individualised treatments, should be given.*	2 2	2 2	0 2
11	Is the needle type (gauge, length, and manufacturer or material) stated?	1 1	0 0	0 0
	Treatment regimen			
12	Are the number or treatment sessions clearly documented? *Where there is variation in the regimen between patients, the mean and range should be reported.*	2 2	2 2	2 2
13	Is the frequency of treatment clearly documented? *Where there is variation in the regimen between patients, the mean and range should be reported.*	2 2	1 1	2 2
	Co-interventions			
14	Are auxiliary techniques clearly documented (e.g. moxibustion, cupping and herbs)?	NA NA	0 0	0 2
15	Are prescribed self-help interventions clearly documented (e.g. exercises and lifestyle advice, such as dietary changes, based on diagnostic criteria)?	NA NA	NA NA	NA NA
	Practitioner background			
16	Practitioner's duration of relevant training stated?	0 0	0 0	0 1
17	Length of clinical experience stated?	0 0	0 0	1 1
18	Expertise in specific condition stated?	0 0	0 0	1 1
	Control intervention)			
19	Is the intended effect of the control intervention and its appropriateness to the research question described? *The control may be an active comparison such as physiotherapy, a minimally active penetrating or non-penetrating sham, or an inert control such as an inactivated TENS machine. Participants may or may not be blinded. In all cases the aims and objectives of the control in relation to the research question should be given.*	2 2	1 1	2 2
20	Are sources that justify the choice of control provided?	1 0	0 0	2 2
21	Are the explanations given to patients of treatment and control interventions stated? *For sham acupuncture this should include the precise wording used, e.g. "another type of acupuncture" or "not acupuncture, but feels similar".*	1 2	0 0	1 1
22	Is the control credible? *If the control treatment is common practice for the condition answer yes. If the control is sham acupuncture or an inert control, answer yes if its credibility has been demonstrated on the target population.*	1 0	2 1	1 1
23	Is the control intervention clearly reported? *A precise description, as for "Needling details" above, and other items if a different regimen from the acupuncture group has been used should be given.*	2 1	2 1	2 1
	Maximum score = 46 minus NA scores	42	44	44
	Independent score	27 25	18 18	32 34
	Score	26 62%	18 41%	33 79%

The candidate's scores are blue (AL), supervisor's black (IR), and agreed scores are in green ink
Key: 2=yes, 1=partially, 0=no, NA=not applicable

Scores differed between raters for intention to treat analysis (ITT), because one (AL) scored all trials zero as none had made specific mention of ITT analysis, and the other (IR) scored the Ma (1999) study full points for ITT on the basis that if attrition was zero, then by default an ITT analysis had been carried out. And to distinguish between studies on quality better, one rater (IR) awarded the Helms (1987) study an additional half mark in recognition of their attention to minimum important differences even if they were, as the author acknowledges, *"arbitrarily defined"*. The Ma (1999) study scored particularly poorly on items related to data collection and analysis methods *[Table 13, items 9, 10, 11, 15, 16]*.

Interventions in trials

The interventions checklist was not piloted before the assessment proper, but there was good agreement between raters on the scores for the Ma (1999) and Thomas (1995) studies with a difference of one and two points respectively *[Table 14, articles 1 (Helms 1987); 2 (Ma 1999); and 4 (Thomas 1999)]*. The difference of six points between raters on the Helms (1987) study was reduced to a difference of two once the information about needling given by Helms (1987) had been teased out to reveal that whilst no depth of insertion or deqi response had been stated, giving a score of zero for items seven and eight, and the precise nature of the manual stimulation had not been described (i.e. even, reducing or reinforcing technique), it was clear that stimulation was manual and one point awarded. Of note, is that the Thomas (1995) study picked up points for providing information about the practitioner's training, experience and expertise not offered by either of the other two. The Thomas (1995) study was awarded an excellent score of 79% for the attention paid to reporting the four different modes of acupuncture and three different modes of transcutaneous electrical nerve stimulation tested. The Helms (1987) study was awarded a good score of 62% for its description of the acupuncture formula and three controls used, and the Ma (1999) study was given a mediocre score of 41% for the description it gave of the three different formulae prescribed according to syndrome patterns in the acupuncture arm of the trial and the hormone treatments in the comparator arm. Thus all three studies had been more rigorous in their administering of the intervention than they had randomisation, for example, with a clear relationship between the better study and intervention assessment scores.

Methods in case-series

The checklist had been piloted on five of the 15 case-series studies and amendments made during its initial development. Following the assessment there were discrepancies in scores of more than three points on nine of the 15 papers with major discrepancies across three of the items in particular: items 10, 13 and 14 in Table 15a and b. Closer inspection revealed problems with interpretation of the question and accompanying guidelines. These differences had led, for example, one rater (AL) to award zero points to studies on account of their unsophisticated and unvalidated practitioner-assessed primary outcomes of *"cured"*, *"much improved"*, *"improved"* and *"no change"*. By contrast the other rater had awarded zero or one point where researchers had sought to validate these outcomes against other measures: changes in prostaglandin levels in menstrual blood (PGF2) in a subgroup of patients treated for dysmenorrhoea (Ge 2000) *[article number 5, Table 15a]*; basal body temperature charts (BBT) and mid-cycle luteinizing hormone (LH) levels in patients treated for menstrual irregularities and amenorrhoea (Tureanu 1994) *[article number 10, Table 15b]*; 'haemological' indices in a subgroup of patients treated for menorrhagia (Wangcheng 1988) *[article number 15, Table 15b]*; and BBT charts and, in a subset, scans to detect follicular development and tests to reveal changes in hormone levels (follicular stimulating hormone, LH, estradiol, progesterone and pituitary prolactin) for ovulatory dysfunction (Mo1993) *[article 19, Table 15b]*. To better facilitate this more sensitive distinction between studies the wording was changed so that studies using measures of *"proven"* validity would be awarded the full two marks and studies using *"credible"* measures would be awarded one mark and studies using neither awarded zero. These changes to item 10 of the questionnaire are presented using the MSWord function track changes in Table 15.

The second item that showed major discrepancies concerned follow-up, where one reviewer (AL) awarded all 15 case-series zero marks according to the guideline *"answer no...where a study does not report the number of patients lost to follow-up,"* and the second reviewer awarded full marks to those who did not report any loss to follow-up, and one mark to those who alluded to but gave an incomplete account of attrition. In order to be sure to award marks for quality criteria attended to, the question was re-worded to ask whether follow-up was *"carefully documented" [item 13, Table 15b]*. This led to all but two of the studies being awarded zero. The study

Table 15a: Scores for methods in non-randomised studies of acupuncture for menstrual conditions – article numbers: 5-9

#	Question	5	6	7	8	9
	Reporting					
1	Is hypothesis or aim clearly described?	1* 0*	0 0	0 0	1 0	0 0
2	In the introduction or discussion, do the authors acknowledge that the study is concerned with and designed only to describe the existing distribution of variables, without regard to causal hypotheses? *(Answer no if authors overstep the data)***	0 0	0 0	0 1	0 0	0 1
3	Are the interventions of interest clearly described? *[Treatments should be clearly described: 2 = score > 74% on Acupuncture Checklist; 1 =score > 50%; & 0 = score < 50%*	0 0	0 0	0 0	1 1	0 0
4	Are the main findings of the study clearly described? *[Simple outcome data (including denominators and numerators) should be reported for all major findings so reader can check major analyses and conclusions]*	2 2	1 1	0 0	1 1	0 0
	External Validity					
5	Is the process used to select the sample completely defined? *(To answer yes, the study must identify the source population and describe how the patients were selected or recruited) (To answer yes the study must give explicit inclusion criteria)*	2 2	1 1	0 0	1 1	0 0
6	Are the characteristics of the patients included in the study clearly described? *(To answer yes, the study must give explicit inclusion criteria)*	1 1	0 0	0 0	1 1	0 0
	Internal Validity – confounding					
7	Were the patients sufficiently homogenous with respect to prognostic risk? Did the study address homogeneity *(To answer yes, the study must describe the distribution of principal confounders such as investigate the prognostic effect of principle confounders age, chronicity and parity)*	2 2	0 0	0 0	0 0	0 0
8	Were patients recruited from the same time-period, so avoiding non-contemporaneous bias through changes in definitions, diagnoses, exposures, disorders and interventions over time? *(Where the time-period over which patients were recruited is not stated the question should be answered no)*	1 0	0 0	0 0	0 0	0 0
9	If sub-group comparisons are made, are the distributions of principal confounders in each group of subjects to be compared clearly described? *(E.g. age, parity, contraception, obesity, severity, chronicity and co-morbidity)*	0 0	NA NA	NA NA	NA NA	NA NA
	Internal Validity – bias					
10	Are the main outcome measures of proven validity accurate (valid and reliable)? *(For studies which refer to other work answer yes. For studies that demonstrate the outcome measures are credible answer partially accurate the question should be answered yes. For studies where outcome measures are clearly described, question should be answered partially)*	1 1	0 0	0 0	0 0	0 0
11	Was an attempt made to blind those measuring the main outcomes of the intervention?	0 0	0 0	0 0	0 0	0 0
12	Was follow-up long enough for important events to occur?	0 1	0 0	0 0	2 2	0 0
13	Was follow-up carefully documented and sufficiently complete? *(This should be answered yes where validity of study is not threatened by a loss that is large in relation to the proportion of patients who did not respond, whose symptoms became worse, or who suffered unacceptable side effects from the intervention. Answer no where losses to follow-up are large or where a study does not report the number of patients lost to follow-up.)*	0 0	0 0	0 0	0 0	0 0
14	Were the statistical tests used to assess the main outcomes appropriate? *(The statistical techniques used must be appropriate to the data. Where little statistical analysis has been undertaken but where there is no evidence of bias, the question should be answered Partially.)*	2 2	NA NA	NA NA	0 0	NA NA
	Maximum score = 28 minus NA scores	28	24	24	26	24
	Independent scores	12 11	2 2	0 1	7 6	0 1
	Score	11.5 41%	2 8%	0.5 2%	6.5 25%	0.5 2%
	Article	*5*	*6*	*7*	*8*	*9*

**The second raters scores are black (IR), the candidate's are blue (AL)and aggregate scores are in green ink.*
***Where agreements involved amendments to the checklist items, the original is shown in red and agreed amendments in green*

References for article numbers 5 to 19: #5 Ge 2000 **6#** Yu 1998 **7#** Jiao 1997 **8#** Wang 1996 **9#** Yong-hai 1995 **10#** Tureanu 1994 **11#** Zhang 1994 **12#** Liu 1992 **13#** Liang 1990 **14#** Zhan 1990 **15#** Wangcheng 1988 **16#** Wang 1987 **17#** Yuquin 1984 **18#** Steinberger 1981 **19#** Mo 1993

Table 15b: Scores for methods in non-randomised studies of acupuncture for menstrual conditions – article numbers: 10-19

#	Article number:	10	11	12	13	14	15	16	17	18	19
	Reporting										
1	Is hypothesis or aim clearly described?	2 1	0 0	0 0	0 0	0 0	2 0	0 0	0 0	1 1	2 0
2	In the introduction or discussion, do the authors acknowledge that the study is concerned with and designed only to describe the existing distribution of variables, without regard to causal hypotheses? ~~*(Answer no if authors overstep the data)*~~	0 0	0 1	0 0	0 0	0 0	0 0	0 0	0 1	0 1	0 0
3	Are the interventions of interest clearly described? *[Treatments should be clearly described: 2 = score > 74% on Acupuncture Checklist; 1 =score > 50%; & 0 = score < 50%)*	1 1	0 0	0 0	0 0	0 0	1 1	0 0	0 0	1 1	0 0
4	Are the main findings of the study clearly described? *[Simple outcome data (including denominators and numerators) should be reported for all major findings so reader can check major analyses and conclusions]*	1 1	1 1	1 1	1 1	1 1	2 2	1 1	1 1	1 1	2 2
	External Validity										
5	Is the process used to select the sample completely defined? ~~*(To answer yes, the study must identify the source population and describe how the patients were selected or recruited)*~~ *(To answer yes, the study must give explicit inclusion criteria)*	1 1	0 0	2 2	0 0	2 2	0 0	2 2	0 0	0 0	2 2
6	Are the characteristics of the patients included in the study clearly described? ~~*(To answer yes, the study must give explicit inclusion criteria)*~~	1 1	0 0	0 0	1 1	0 0	2 2	2 2	0 0	0 0	1 1
	Internal Validity – confounding										
7	~~Were the patients sufficiently~~ Did the study address homogeneity? ~~homogenous with respect to prognostic risk?~~ *(To answer yes, the study must investigate the prognostic effect ~~describe the distribution~~ of principal confounders such as age, chronicity and parity)*	0 0	0 0	2 2	2 2	1 1	0 0	0 0	1 1	0 0	0 0
8	Were patients recruited from the same time-period, so avoiding non-contemporaneous bias through changes in definitions, diagnoses, exposures, disorders and interventions over time? *(Where the time-period over which patients were recruited is not stated the question should be answered no)*	0 0	0 0	1 0	0 0	0 0	0 0	1 0	0 0	2 1	2 2
9	If sub-group comparisons are made, are the distributions of principal confounders in each group of subjects to be compared clearly described? *(E.g. age, parity, contraception, obesity, severity, chronicity and co-morbidity)*	NA NA	NA NA	NA NA	NA NA	NA NA	0 0	NA NA	NA NA	NA NA	0 0
	Internal Validity – bias										
10	Are the main outcome measures of proven validity ~~accurate (valid and reliable)~~? *(For studies which refer to other work ~~answer~~ Yes. For studies that demonstrate the outcome measures are ~~accurate~~ credible answer partially ~~the question should be answered yes. For studies where outcome measures are clearly described, the question should be answered partially)~~*	1 1	0 0	0 0	0 0	0 0	1 1	0 0	1 1	0 0	1 1
11	Was an attempt made to blind those measuring the main outcomes of the intervention?	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0
12	Was follow-up long enough for important events to occur?	1 2	1 2	1 0	0 1	2 2	0 0	1 2	0 1	1 2	0 0
13	Was follow-up carefully documented and sufficiently complete? *(This should be answered yes where validity of study is not threatened by a loss that is large in relation to the proportion of patients who did not respond, whose symptoms became worse, or who suffered unacceptable side effects from the intervention. Answer no where losses to follow-up are large or where a study does not report the number of patients lost to follow-up.)*	0 0	0 0	0 0	0 0	0 0	1 1	0 0	0 0	2 2	0 0
14	Were the statistical tests used to assess the main outcomes appropriate? ~~*(The statistical techniques used must be appropriate to the data. Where little statistical analysis has been undertaken but where there is no evidence of bias, the question should be answered Partially.)*~~	NA NA	NA NA	0 0	0 0	0 0	1 1	NA NA	NA NA	NA NA	1 1
	Maximum score = 28 minus NA scores	24	24	26	26	26	28	24	24	24	28
	Independent scores	8 8	2 4	7 5	4 5	6 6	10 8	7 7	3 5	8 9	11 9
	Score	8 33 %	3 13 %	6 23 %	4.5 16 %	6 23 %	9 32 %	7 29 %	4 17 %	8.5 35 %	10 36 %
	Article	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*	*19*

**The second rater's scores are black (IR), the candidate's are blue (AL) and aggregate scores are in green ink.*
***Where agreements involved amendments to the checklist items, these are given in red using MSWord 'track changes'*

KEY: 2 =yes; 1=partially; 0=no; NA=not applicable

by Steinberger (1981) was awarded one point in recognition of its honesty and a follow-up rate of 92% [*article number 18, Table 15b]*, and the study by Wangcheng (1988) was awarded one point for admitting that only 20 (67%) of its 30 patients were followed for three full menstrual cycles [*article number 15, Table 15b]*.

The need to record as *"not applicable"* the use of appropriate statistical analysis was also noted at this point of the agreement process, as one rater (AL) had awarded full marks to studies that had not used statistical tests, so overstepping the guideline *"where little statistical analysis has been undertaken but where there is no evidence of bias, the question should be answered partially,"* and one mark to a more sophisticated study that had used tests, even if imperfectly (Ge 2000). This linking of two subtly distinct concepts of bias and analysis was considered the problem and it was decided that the question be restricted to analysis. The subsequent amendment is recorded in Table 15, item 14, and led eight of the case-series to be considered inference free papers not in need of statistical analysis and excluded from this criteria. Of the remaining seven, the study by Ge (2000) was awarded one mark for reporting two chi-squared tests explicitly, if not perfectly, and one t-test implicitly but correctly *[article 5, Table 15a]*; the study by Mo (1993) was awarded one mark for reporting many t-tests implicitly but correctly *[article 19, Table 15b]*; and one mark was awarded to the study by Wangcheng (1988) for similar reasons. Zero points were awarded to the four studies that claimed to draw statistical inferences without the necessary statistical tests (Wang 1996; Liu 1992; Liang 1990; Zhan 1990) *[article numbers 8, 12, 13,14 Table 15a & b]* - and perhaps highlight a plausible case for a 'minus one' category distinction.

These changes brought the assessment scores closer together but four items still showed major discrepancies: four, five, six and seven in Table 16. On item four and the clarity of findings, one rater (AL) awarded all studies that had not used statistical analyses and *clearly* stated the proportion of patients that were *"cured"* etc, the full two points, and those that had presented incomplete statistical data one point. To reward the more sophisticated of the studies, scores were revised so that Ge (2000) was awarded two points for three clear tables of analyses *[article 5, Table 15a]*; Wangcheng (1988) for two clear Tables *[article 15, Table 15b]*; Mo (1993) for three clear Tables *[article 19, Table 15b]*; but zero to Jiao 1997 *[article 7, Table 15a]*, who

merged criteria and results; and Yong-hai (1995) *[article 9, Table 15a]*, who did not even define a cure.

Item five asked whether the process used to select the sample had been completely defined whilst six asked about the sample's characteristics. Consequently, in order to award studies for a detailed description of the sample's characteristics and not confuse this with the studies inclusion criteria, the guideline *"the study must give explicit inclusion criteria"* was moved to qualify item five better and the process used to select the sample. These amendments improved the level of rater agreement and are shown with the MSWord function track changes in Tables 15a & b.

Finally, to distinguish question seven about homogeneity from question six and the samples characteristics, this question was changed from asking for a *"description of the distribution of confounders"*, to ask whether the study *"investigated the prognostic effect of confounders"* such as age, chronicity and parity. And for the purposes of this review, to improve sensitivity and use the scale to full effect, the reviewers awarded two points where a study had looked at two confounders (Ge 2000; Liu 1992; Liang 1990) *[article numbers 5, 12 and 13, Table 15a & b]*, and one point where it had looked at one (Zhan 1990; Wang 1987*)[article numbers 14 and 17, Table 15b]*.

These revisions led to a convergence in scores across the fifteen studies *[Tables 15a and b]*. Scores ranged from a very poor 2% (Jiao 1997; Yong-hai 1995 – articles 7 and 9) to a mediocre 41% (Ge 2000, article 5), with six of the 15 studies scoring below 20% (Yu 1998; Jiao 1997; Yong-hai 1995; Zhang 1994; Liang 1990; Yuquin 1984), four of the fifteen scoring between 20 an 30% (Wang 1996; Liu 1992; Zhan 1990; Wang 1987), and the remaining five scoring from 30 to 41% (Ge 2000; Tureanu 1994; Wangcheng 1988; Steinberger 1981; Mo 1993) *[Table 17 – summary of scores]*.

<u>Interventions in case-series</u>

The interventions checklist, when applied to the case-series studies, revealed convergence between raters with less than 3 points difference in scores for 11 of the 15 studies, but for the remaining four there was as many as 11 points difference.

Table 16a: Scores for interventions in studies of acupuncture for menstrual conditions – article numbers: 5 to 9

Item	Article number:	5	6	7	8	9
	Acupuncture rationale					
1	Is the style of acupuncture used defined? *The approach taken should be stated, e.g. traditional Chinese medicine or western medical.*	2 2	1 0	2 2	2 1	0 1
2	Is the rationale for treatment and whether it is individualised or standardised given? For *example according to syndrome patterns, segmental levels or trigger points should be ~~stated~~ defined. (Answer partially if merely stated, and yes if theoretical and diagnostic criteria are described)*	1 1	0 0	1 1	2 2	1 1
3	Are literature sources to justify the rationale provided?	0 0	0 0	0 0	0 0	0 0
	Needling details					
4	Are points used described with standard nomenclature (e.g. ST36, Zusanli) or in terms of anatomical location?	2 2	2 2	2 2	2 2	2 2
5	Are specific point locations, and whether unilateral or bilateral, given for a formula of points? Or, where treatment is individualised, are typical points and the range of points used given?	1 1	2 1	1 1	1 2	2 2
6	Is the number of needle insertions specified? *A simple total of needle insertions for a formula of points, or as a mean and range where the number of needles varies between patients - aim to answer yes or no to this question unless number is stated but discrepancies exist*	0 0	0 0	0 0	0 0	0 0
7	Are depths of insertion given (e.g. *cun* or tissue level) and whether standardised or individualised?	2 2	0 0	0 0	2 2	0 0
8	Are responses elicited described (e.g. deqi or twitch response)? *(Answer yes if sensations described (e.g. ache, tingling, heat, distension or comfortable muscle twitch) and partially if induction of a needle sensation (degi or twitch) merely stated)*	1 1	1 1	1 1	2 2	0 0
9	Is needle stimulation described (e.g. manual or electrical)? *(Answer partially if method merely stated, and yes if technique is given e.g. "reducing" or "reinforcing" or Hz applied)*	2 2	2 2	1 1	2 2	2 2
10	Is needle retention time specified? *The standard time needles were retained, or the mean and range for individualised treatments, should be given.*	2 2	2 2	2 2	2 2	2 2
11	Is the needle type (gauge, length, and manufacturer or material) stated?	2 1	0 0	0 0	1 1	0 0
	Treatment regimen					
12	Are the number or treatment sessions clearly documented? *Where there is variation in the regimen between patients, the mean and range should be reported.*	1 1	0 0	2 2	1 1	1 1
13	Is the frequency of treatment clearly documented? *Where there is variation in the regimen between patients, the mean and range should be reported.*	2 2	1 1	2 2	1 1	1 1
	Co-interventions					
14	Are auxiliary techniques clearly documented (e.g. moxibustion, cupping and herbs)?	2 1.5	0 0	NA NA	2 1	NA NA
15	Are prescribed self-help interventions clearly documented (e.g. exercises and lifestyle advice, such as dietary changes, based on diagnostic criteria)?	NA NA	NA NA	1 1	2 1	NA NA
	Practitioner background					
16	Practitioner's duration of relevant training stated?	0 0	0 0	0 0	0 0	0 0
17	Length of clinical experience stated?	0 0	0 0	0 0	1 0	0 0
18	Expertise in specific condition stated?	0 0	0 0	0 0	0 0	0 0
	Control intervention)					
19	Is the intended effect of the control intervention and appropriateness to research question described?	NA	NA	NA	NA	NA
20	Are sources that justify the choice of control provided?	NA	NA	NA	NA	NA
21	Are the explanations given to patients of treatment and control interventions stated	NA	NA	NA	NA	NA
22	Is the control credible?	NA	NA	NA	NA	NA
23	Is the control intervention clearly reported?	NA	NA	NA	NA	NA
	Maximum score = 46 minus NA scores	34	34	34	36	32
	Independent scores	20 18.5	11 9	15 15	23 20	11 12
	Score	19.3 57%	10 29%	15 44%	21.5 60%	11.5 36%
Article		*5*	*6*	*7*	*8*	*9*

**The candidate's scores are blue (AL), supervisor's black (IR), and agreed scores are in green ink*
KEY: 2 =yes; 1=partially; 0=no; NA=not applicable

References for article numbers 5 to 19: #5 Ge 2000 **6#** Yu 1998 **7#** Jiao 1997 **8#** Wang 1996 **9#** Yong-hai 1995 **10#** Tureanu 1994 **11#** Zhang 1994 **12#** Liu 1992 **13#** Liang 1990 **14#** Zhan 1990 **15#** Wangcheng 1988 **16#** Wang 1987 **17#** Yuquin 1984 **18#** Steinberger 1981 **19#** Mo 1993

Table 16b: Scores for interventions in un-controlled studies of acupuncture for menstrual conditions – article numbers: 10 to 19

Item	Article number:	10	11	12	13	14	15	16	17	18	19
	Acupuncture rationale										
1	Is the style of acupuncture used defined? *The approach taken should be stated, e.g. traditional Chinese medicine or western medical.*	2 2	2 1	2 2	1 1	1 2	2 2	1 1	1 0	2 2	2 2
2	Is the rationale for treatment and whether it is individualised or standardised given? *The rationale for prescribing treatment, for example according to syndrome patterns, segmental levels or trigger points should be ~~stated~~ defined. (Answer partially if merely stated, and yes if theoretical and diagnostic criteria are described)*	2 2	2 2	2 2	0 0	2 2	2 2	2 2	2 2	2 2	2 2
3	Are literature sources to justify the rationale provided?	2 2	0 0	0 0	0 0	0 0	1 0	0 0	0 0	2 1	0 0
	Needling details										
4	Are points used described with standard nomenclature (e.g. ST36, Zusanli) or in terms of anatomical location?	2 2	1 2	1 2	2 2	2 2	2 2	2 2	2 2	2 2	2 2
5	Are specific point locations, and whether **unilateral or bilateral**, given for a formula of points? Or, where treatment is individualised, are typical points and the range of points used given?	1 2	1 2	1 1	1 1	1 1	1 1	1 1	1 1	2 2	1 0
6	Is the number of needle insertions specified? *A simple total of needle insertions for a formula of points, or as a mean and range where the number of needles varies between patients - aim to answer yes or no to this question unless number is stated but discrepancies exist*	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0
7	Are depths of insertion given (e.g. *cun* or tissue level) and whether standardised or individualised?	0 0	0 0	0 0	0 1	0 0	0 0	0 0	0 0	0 0	0 0
8	Are responses elicited described (e.g. deqi or twitch response)? *(Answer yes if sensations described (e.g. ache, tingling, heat, distension or comfortable muscle twitch) and partially if induction of a needle sensation merely stated)*	0 0	2 2	0 0	0 0	1 1	1 1	0 0	0 0	1 1	2 2
9	Is needle stimulation described (e.g. manual or electrical)? *(Answer yes if technique described eg. "reducing" or "reinforcing" technique stated or Hz applied, partially if method merely stated)*	0 0	2 2	1 1	2 2	2 2	2 2	1.5 1.5	2 2	2 2	2 2
10	Is needle retention time specified? *The standard time needles were retained, or the mean and range for individualised treatments, should be given.*	2 2	0 0	2 2	0 0	2 2	2 2	2 2	2 2	2 2	2 2
11	Is the needle type (gauge, length, and manufacturer or material) stated?	2 2	0 0	0 0	1 1	0 0	1 0	0 0	0 0	2 2	0 0
	Treatment regimen										
12	Are the number or treatment sessions clearly documented? *Where there is variation in the regimen between patients, the mean and range should be reported.*	1 1	1 1	1 1	1 1	2 2	1 1	1 1	2 2	2 2	1 1
13	Is the frequency of treatment clearly documented? *Where there is variation in the regimen between patients, the mean and range should be reported.*	2 2	1 1	1 1	1 1	2 2	1 1	1 1	0 0	2 2	1 1
	Co-interventions										
14	Are auxiliary techniques clearly documented (e.g. moxibustion, cupping and herbs)?	NA NA	NA NA	1 1	NA NA	1 1	NA NA	1 1	1 1	2 1	NA NA
15	Are prescribed self-help interventions clearly documented (e.g. exercises and lifestyle advice, such as dietary changes, based on diagnostic criteria)?	NA NA	NA NA	NA NA	NA NA	NA NA	NA NA	NA NA	0 1	NA NA	NA NA
	Practitioner background										
16	Practitioner's duration of relevant training stated?	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0
17	Length of clinical experience stated?	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0
18	Expertise in specific condition stated?	1 1	0 0	0 0	0 0	0 0	0 0	0 0	0 0	1 1	0 0
	Control intervention items 19 to 23 = NA = 10										
	Maximum score = 46 -- NA scores	32	32	34	32	34	32	34	36	34	32
	Independent scores	17 18	12 13	12 13	9 10	16 17	16 14	12.5 12.5	13 13	23 22	15 14
	Score	17.5 55%	12.5 39%	12.5 37%	9.5 30%	16.5 49%	15 47%	12.5 37%	13 36%	22.5 66%	14.5 45%
	Article	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*	*19*

*The candidate's scores are blue (AL), supervisor's black (IR), and agreed scores are in green ink

Key *2= yes 1= partially 0= no NA = not applicable*

Closer inspection found good correlation between raters on 17 of the 23 items in the checklist with major discrepancies on items two, six, eight, nine, 12 and 13 *[Table 16]*. Item two relates to the treatment rationale; item six the specific number of needle insertions; item eight the needle response elicited; item nine needle stimulation; item 12 the number of treatment sessions prescribed and 13 the frequency of treatments.

The lack of convergence on such seemingly simple criteria as in item six – *"is the number of needle insertions specified?"* – occurred because neither rater applied the criteria strictly and instead would give a point if able to deduce the approximate number of insertions from information about the points used. However, as item five awards points for information about the points used, it was agreed the criteria in item six be applied rigidly to better award this additional information and the guideline amended to read, *"- aim to answer yes or no to this question unless number is stated but discrepancies exist"*. This same problem also occurred for items 12 and 13 where the criteria *"a mean and range should be reported where there is variation in regimen between patients"*, was later better applied to distinguish between studies so that where the course of treatment was stated (e.g. five daily sessions) but the mean and range not given despite the case study used to illustrate treatment having had two or more courses of five daily sessions, one point was awarded – indeed, the great majority of the 15 case series failed to adequately define the course of treatment patients received. The exceptions were awarded two points: Jiao 1997, Zhan 1990, Yuquin 1984 and Steinberger 1981 *[articles 7, 14, 17 and 18 in Tables 16a and b]*.

The other items where convergence was poor required better guidance for raters. For item two about the acupuncture treatment rationale, the guidance was amended to read, *"The rationale for prescribing treatment, for example according to syndrome patterns, segmental levels or trigger points should be defined. (Answer partially if merely stated, and yes if theoretical and diagnostic criteria are described)."* This meant zero points were awarded to Liang (1990) who provided no rationale for the study and acupuncture at all *[article 13, Table 16b]*; one point was awarded to Ge (2000) and Jiao (1997) who merely stated the TCM diagnostic criteria for treatment - e.g. "excess" or "deficient syndrome" *[article numbers 5 and 7, Table 16a]*; and two points were awarded to the rest of the studies which provided a detailed rationale for

treatment and defined their diagnostic criteria (Yu 1998; Wang 1996; Yong-hai 1995; Tureanu 1994; Zhang 1994; Liu 1992 Zhan 1990; Wangcheng 1988; Wang 1987; Yuquin 1984; Steinberger 1988; Mo 1993).

Item eight about needle sensations was refined to better recognise the information provided so that two points were awarded to Mo (1993), Zhang (1994) and Wang (1996) *[articles 19, 11 and 16, Table 16b],* who described the nature of the sensation or response elicited (e.g. aching, distension, heat), and one point for those that merely stated that a needle sensation or deqi was induced (Ge 2000; Yu 1998; Jiao 1997; Zhan 1990; Wangcheng 1988; Steinberger 1988) *[articles 5, 6, 7, 14, 15 & 18 tables 16a and b].* Item nine that asked about needle stimulation was elaborated upon to better use the scores so that ten of the studies were awarded two points for describing the method *and* technique used to stimulate the needles, such as *"uniformly reinforcing and reducing method"* or *"very gently manipulated until sensation felt"* (Ge 2000; Yu 1998; Wang 1996; Yonghai 1995; Zhang 1994; Liang 1990; Zhan 1990; Wangcheng 1988; Wang 1987; Yuquin 1984; Steinberger 1981; Mo 1993) *[articles 5, 6, 8, 9, 11,13, 14, 15, 16, 17 & 18, Tables 16a and b].* Jiao (1997) was awarded one point for stating manual stimulation *[article 7, Table 16a]*, and Tureanu (1994) zero for no mention of needle technique *[article 10, Table 16b].*

The amendments and agreements reached for items two, six, eight, nine, 12 and 13 brought 13 of the 15 scores within 3 points of each other. Discrepancies of 4 points remained on two of the papers, which were resolved after a further assessment to bring all 15 within 3 points. The subsequent scores ranged from a poor assessment score of 29% to a good score of 66%, with a mean mediocre score of 44% *[Table 17].* Again those with better scores for the intervention had better corresponding scores for methodology, and as with the trials, more attention had been paid to the acupuncture intervention than the study design and methods.

3.2.3 Characteristics of the included studies

Between the years 1981 to 2000, 15 case-series and three controlled trials of acupuncture in the treatment of menstrual disorders were published and all found in favour of acupuncture (Ge 2000; Ma 1999; Yu 1998; Jiao 1997; Wang 1996; Thomas

Table 17: Summary of checklist scores – values are percent (denominator)

Article number	Author	QUALITY OF THE METHODOLOGY	QUALITY OF THE INTERVENTIONS
TRIALS		*Max score = 32*	*Max score = 46*
1	Helms 1987	51% (32)	62% (42)
2	Ma 1999	23% (32)	41% (44)
4	Thomas 1995	55% (32)	79% (42)
	RANGE (mean)	23 to 55% (43%)	41 to 79% (61%)
CASE SERIES		*Max score = 28*	*Max score = 36*
5	Ge 2000	41% (28)	57% (34)
6	Yu 1998	8% (24)	29% (34)
7	Jiao 1997	2% (24)	44% (34)
8	Wang 1996	25% (26)	60% (36)
9	Yong-hai 1995	2% (24)	36% (32)
10	Tureanu 1994	33% (24)	55% (32)
11	Zhang 1994	13% (24)	39% (32)
12	Liu 1992	23% (26)	37% (34)
13	Liang 1990	16% (26)	30% (34)
14	Zhan 1990	23% (26)	49% (32)
15	Wangcheng 1988	32% (28)	47% (34)
16	Wang 1987	29% (24)	37% (32)
17	Yuquin 1984	17% (24)	36% (34)
18	Steinberger 1981	35% (24)	66% (36)
19	Mo 1993	36% (28)	45% (34)
	RANGE (mean)	2 to 41% (23%)	29 to 66% (44%)

1995; Yong-hai 1995; Tureanu 1994; Zhang 1994; Mo 1993; Liu 1992; Liang 1990; Zhan 1990; Wangcheng 1988; Helms 1987; Wang 1987; Yuquin 1984; Steinberger 1981). A summary of scores awarded for both methodological quality and quality of the interventions in this review is given in Table 18. The 15 case series had in common a number of serious methodological weaknesses: failure to describe the way the cases were obtained; incomplete descriptions of the intervention; and the use of unreliable and unsophisticated primary measures of outcome (Ge 2000; Yu 1998; Jiao 1997; Wang 1996; Yong-hai 1995; Zhang 1994; Mo 1993; Liu 1992; Liang 1990; Zhan 1990; Wangcheng 1988; Wang 1987; Yuquin 1984). An attempt to validate practitioner assessments according to criteria such as cured/effective/no effect against more credible measures of change, such as BBT charts, changes in hormone levels, and scans to detect follicular development, were a positive feature in four of the studies (Ge 2000; Tureanu 1994; Mo 1993; Wangcheng 1988), and it is notable that these were four of the five to score better overall for their methods and intervention quality, the fifth being the study by Steinberger (1981). The study by Ge (2000) scored 41% for methods and 57% for intervention; Tureanu (1994) scored 33% for methods and 55% for intervention; Mo (1993) scored 36% for methods and 45% for

intervention; Wangcheng (1988) scored 32% for methods and 47% for intervention and Steinberger (1981) scored 35% for methods and 66% for intervention. Thus these five studies demonstrated greater rigour than the rest when defining the rationale for acupuncture and the precise treatment regimen - but the data collection and analysis methods were again poorly defined.

A 'one-size-fits-all' formula of points was used in three of these five case-series, of which two were Chinese and one European (Steinberger 1981): Ge (2000) prescribed SP6 REN4 BL26 for primary dysmenorrhoea in women aged 17 to 19 years; Mo (1993) prescribed BL18 BL23 REN3 REN4 SP6 EXTRA19 for anovulation in patients aged 21 to 39 years; and Steinberger (1981) prescribed ST36 SP6 LI4 REN4 to women aged 19 to 42 with primary dysmenorrhoea. By contrast, one Chinese (Wangcheng 1988) and one European study (Tureanu 1994) administered individualised treatment according to differential syndrome diagnoses. By paying particular attention to underlying TCM theory and diagnostic criteria, and by listing the typical points used for the different syndrome diagnoses, they succeed in providing a fairly clear description of the individualised TCM acupuncture employed in the treatment of 30 women with menorrhagia (Wangcheng 1988) and 38 women with post oral contraceptive menstrual irregularities and amenorrhoea (Tureanu 1994). And were there also precise details about the number of treatments provided, or mean and range, along with the practitioner's training and expertise, and needling sensations elicited, these studies would have achieved "excellent" scores for the intervention as opposed to mediocre to good scores. However these studies are notable, as they meet the standards for authentic and credible treatment, which members of the British Acupuncture Council can be expected to achieve (Birch & Felt 1999; Maciocia 1998; Lyttleton 2004).

Of the three controlled trials, the American (Helms 1987) and Swedish (Thomas 1995) studies demonstrate greater rigour. The Chinese study (Ma 1999) involved a most respectable 215 participants, but did not define the recruitment procedures employed; failed to randomise participants to groups; pays no respect to prognostic risk; gave only a partial rationale for, and incomplete descriptions of, the experimental (acupuncture) and comparator (drug therapy) interventions; used unreliable and unsophisticated measures of global change and failed to give a rationale for, or define

the data collection and analysis methods for, the potentially more credible measure of hormone levels. The study by Thomas (1995) also failed to randomise, whilst the study by Helms (1987) used an inadequate method. Neither trials' provide sufficient evidence of baseline comparability; both use inappropriate statistical tests; and fail to acknowledge the risk of erroneous positive results (type 1 error) owing to small sample size. Thomas (1995) also fails to acknowledge the potential bias from "carry-over-effects" in a crossover design, despite evidence that improvements in the group receiving active and sham acupuncture were cumulative and subgroup analyses therefore invalid. However, these two trials did acknowledge the risk from prognostic variables; provided details of those lost to follow-up; better defined the interventions and controls; used more credible patient assessed outcomes; and more clearly described their data collection and analysis methods.

4. Discussion

4.1 The current evidence base for acupuncture in the treatment of menorrhagia and related idiopathic menstrual disorders primary dysmenorrhoea and irregular menstruation

Where the quantity and quality of evidence allows, the principle aim of a systematic review is to combine like with like in order to synthesise relevant and reliable outcomes as to the effectiveness of a specific intervention upon a defined population, evaluated according to rigorous methods, using standard validated health assessment measures (Clarke & Oxman 2000; White 2002). However, where there is a lack of studies, the objective of a review necessarily becomes more modest, aiming merely to provide an overview of the range, type and quality of evidence that currently exists in order to establish the foundations upon which future research can be built (Goodman 1993; Light & Pillemer 1984). The scarcity of previous research to evaluate the effectiveness of acupuncture in the treatment of menorrhagia, even when methodological criteria was widened beyond the controlled trial to include case-series studies, determined that this review also consider findings from studies of acupuncture in the treatment of the commonly associated idiopathic menstrual disorders primary dysmenorrhea and irregular menstruation (Ruta et al 1995). And to further assure the comprehensiveness of the review findings, acupuncture studies of other gynaecological conditions such as amenorrhoea were included where changes in

volume of menstrual loss, cycle regularity, and pain at menstruation were outcomes of interest.

This review generated a total of 18 studies, all of which found in favour of acupuncture. Of these, 15 were case-series studies and three were controlled trials. Just two of the studies in the review, both of which were case-series, focused upon the treatment of menorrhagia (Zhang 1994; Wangcheng 1988). In line with the current consensus (Birch & Felt 1999; NHS CRD 2001), all 15 of the studies from China were found to be methodologically poor and to demonstrate a considerable lack of critical awareness, with not one author considering that practitioner assessment according to criteria such as *"cured", "effective" and "no effect"* might be inaccurate and unreliable as a measure of change; although attempting to validate such outcomes against more credible measures was a positive feature of the trial conducted in China (Ma 1999) and four of the case-series (Ge 2000; Tureanu 1994; Mo 1993; Wangcheng 1988). Only one of the Chinese studies (Wangcheng 1988) reported on, or gave any indication of attrition, with most simply stating the number of patients involved in the study and implying 100% compliance to treatment and follow-up. Moreover, the lack of rigour in the reporting of methodology was evident in the reporting of the acupuncture treatment protocol, meaning that the authenticity or credibility of treatments is often difficult to determine. These studies therefore serve only to promote interest, and with sample sizes ranging from a respectable 30 to an impressive 225, highlight the opportunities that were missed to generate compelling evidence. In particular, with advice on methodology the one comparative study amongst the case series from China (Ma 1999), with a sample size over 200, could have provided strong evidence of effectiveness for electro-acupuncture formulae according to syndrome patterns versus hormone treatments in the treatment of post-medication amenorrhoea and led the field.

Thus the most promising findings come from the trials conducted in America (Helms 1987) and Sweden (Thomas 1995). These pilot studies were carried out with some rigour in acupuncture and control interventions, and with equal attention to randomisation procedures and the numbers needed to begin to make statistical inferences, could have added meaningfully to the evidence base for primary dysmenorrhoea. For example, if the 48 patients in the Helms (1987) study had been

randomised to two as opposed to four groups, the study power and its ability to detect statistically significant differences between the acupuncture and control intervention would have been enhanced. The study by Thomas (1995) involved an acupuncture group receiving one of four different modes of acupuncture, including sham acupuncture, prior to five consecutive menstrual cycles; and a TENS group receiving one of three different modes of TENS, including placebo TENS, prior to four consecutive menstrual cycles. This parallel within group crossover design may have benefited from a "washout" period of one month between each mode of acupuncture to avoid the contrast between the three different modes of active acupuncture and sham acupuncture being contaminated by a residual or lag effect from treatment (Grimm 2002) – although given the sample size of just 19 in the acupuncture and 12 in the TENS group, a simple comparison of an active acupuncture intervention with an active TENS intervention would have elicited more valid results, with this most complex design considered where a large trial is justified with sufficient power to detect an effect between treatment modes. Again, with advice on design and methods, these shortcomings might have been avoided and more compelling evidence generated. But for these studies the CONSORT statement (Begg 1996), which has been responsible for a considerable improvement not only in the quality of trial reports, but also their design and conduct, came too late (Altman et al 2001; Moher et al 2001b).

4.2 Strengths and weaknesses of the review

This systematic review represents the first attempt to synthesise evidence for the effectiveness of acupuncture in the treatment of menorrhagia. It was conducted in accordance with the guidelines for carrying out systematic reviews recommended by the NHS Centre for Reviews and Dissemination in 1996 and 2001, and employs many of the review methodologies advised to minimise bias including: an extensive search strategy; the use of checklists to assess methodological quality; and by extension, a checklist to assess the quality of interventions in studies of acupuncture; a data extraction sheet to facilitate study selection; and two reviewers to assess study selection and to apply the quality criteria. All these features serve to improve the precision of the review.

It is worth noting however that the checklists used were developed for the purposes of this review and have not been validated; that is they have not undergone the programme of tests that would be required to be sure they are reliable in their assessments. A possible set of experiments to achieve such construct validity might begin by using the extreme groups design where the checklist is tested to see whether or not it is able to distinguish between studies that are either very good or very bad. Once this is established the ultimate test of usefulness would be its ability to make much finer discriminations amongst studies of similar methodological quality. Clearly, validating a scoring system for reviews of acupuncture studies will become a more pressing concern as quality improves.

The review did not search the grey literature, for example conference proceedings, or search relevant journals before 1985, the start of AMED. It also did not contact authors for further information about the studies, or translate the 14 foreign language papers elicited by the search strategy and given in Table 18.

Table 18: Foreign language papers

Electronic database	Scope of search	References	Potentially relevant studies
AMED	1985 to May 2003	46	2
MEDLINE	1966 to April week 4 2003	193	7
EMBASE	1980 to 2003 week 19	79	7
CINAHL	1982 to May week 1 2002	1	0
	TOTAL	**319**	**16**
	Excluding duplicates (n=2)		**14**
	Chinese (n=2) German (n=4) Italian (n=4) Bulgarian (1) Russian (1) Ukranian(1) Serbo-Croatian (1)		

4.3 Recommendations for future research

In the absence of any credible evidence of effectiveness for acupuncture in the treatment of menstrual disorders, there is a real opportunity for acupuncture students, practitioners and colleges to engage in research that can add in a meaningful way to the evidence base and be accomplished with relatively few resources beyond those available in clinical practice or perhaps some specialist advice (Thomas & Fitter 2002). This would include the N=1 design, uncontrolled before-and-after study and case-series *[Figure 3]* - because a prospective case-series that has been planned and

conducted carefully to ensure that the data collected on naturally occurring eligible patients in a practice setting is uniformly reliable and complete is of value and has yet to be conducted. Qualitative investigations to understand patients' treatment experiences and their preferences for treatment would also add much to this evidence base. All such studies would provide the groundwork for more rigorous exploratory trials to determine the rationale for a definitive trial.

The exploratory trial, nested qualitative study and complementary survey presented in Chapters Three, Four and Five of the thesis were undertaken with the explicit aim of making such a substantive contribution to research in this field.

5. Conclusions

This systematic review of acupuncture for menorrhagia found just two studies that had sought to evaluate acupuncture in the treatment of menorrhagia or excessive menstrual loss. One evaluated the treatment of 50 (Zhang 1994) and the other 30 women (Wangcheng 1988). The positive findings from these studies are purely suggestive as both were case series studies, and in addition to the bias and imprecision inherent in the study design, both suffered from a lack of methodological rigour whereby recruitment procedures and interventions were poorly defined and validated or credible primary outcomes were lacking. The majority of the evidence available concerned the treatment of primary dysmenorrhoea (12 of 18 studies), and most were case series studies (10/12); but two pilot trials indicate a possible beneficial effect from acupuncture (Helms 1987; Thomas 1985), although a lack of methodological rigour, such as inadequate randomisation procedures and inappropriate statistical tests, small sample sizes and evidence of confounding from carry-over effects diminish the potential usefulness of results. Well-conducted exploratory studies – from case-series to randomised trials - would usefully add to this evidence base.

ANNEX A: Acupuncture for Menorrhagia Descriptive Systematic Review: Table of included studies

Study	Design and Methods	Participants	Interventions	Outcomes	Findings	Notes
Helms 1987 Primary dysmenorrhoea #1	RCT Parallel design ***Study QA score: 51% (16.3/32)*** Justification for sample size: no Randomisation secure: no - random numbers table Recruitment procedures: gynaecologist referral, informed consent Recruitment time-period: not stated Number of participants: 48 Number in each group: Real acupuncture (RA) 11 Placebo acupuncture (PA) 11 Standard control (SC) 11 Visitation control (VC) 10 Allocation concealment adequate: no - practitioner researcher allocated participants to groups Blinding: participants only aware if randomised to "control" or "acupuncture" categories Assessor blinding: independent but not blind Baseline comparability: no Lost to follow-up: 5 2 pregnancies, 2 job relocation, 1 lack of interest. 2 from the visitation control group and 1 each from the other 3 groups Intention to treat: no Minimum important differences: yes, but "arbitrary" Appropriate statistical analyses: no – groups too small for analysis of variance, actual probability values not given, few absolute numbers Alternative analyses possible: partial	Location: USA Source population: not stated Inclusion criteria: gynaecologist diagnosis of primary dysmenorrhoea according to "classical criteria" - recurrent pain within 2 years of regular menses, absence of pelvic pathology. Exclusion criteria: not stated Characteristics of participants at intervention commencement: *Age*: average 28 years. RA group mean 26.3 (Standard Error 1.53), PA group mean 30 (SE 1.68), SC group mean 31 (SE 2.37), VC group mean 25.3 (SE 1.5). *Ethnicity*: not stated *Socio-economic group*: not stated *Contraception*: not stated *Parity:* RA group 2/11, PA group 3/11, SC group 6/11, VC group 1/10. *Obesity*: not stated *Severity*: given as 6 months pre-treatment pain scores: RA group 147.09 (SE 38.16), PA group 155.46 (SE 40.51), SC group 97.91 (SE 32.27), VC group . *Chronicity*: given as age at onset: RA group 14.6 (SE 1.68), PA group 15.5 (SE 1.71), SC group 15.5 (SE 1.41), VC group 17.1 (SE 1.54).	Acupuncture formula compared to 3 control interventions ***Interventions QA score: 62% (26/42)*** Setting: Gynaecology Clinic, Oakland, CA Practitioner details: not stated Style of acupuncture: not stated Rationale: TCM protocol described by Mussat 1980 Points: 12 insertions at REN2, REN4 and bilaterally at SP4, KID3, ST36, ST30 and P6. Point locations: verified with a point detector measuring decreased skin electrical resistance at points. Needling details: manual stimulation, needle type, depth of insertion and response not stated Number of treatments: implicit Needle retention time: 30 – 40 minutes (until usual erythema blanched) Frequency of treatments: 1 weekly for 3 weeks a month (to avoid menses), for 3 months. Co-interventions: none Control intervention: (1) *"Placebo acupuncture"* - 12 insertions on lateral thighs and arms at non-meridian points as verified by point detector measuring skin electrical resistance. Treatment regimen (number of treatments, frequency & needle retention time) as for real acupuncture. (2) *Standard control*: initial visit only. Continued using NSAID's and COCP *(3) Visitation control*: visited physician once a month for 3 cycles in addition to usual NSAID's and COCP. Rationale/sources that justify control: partial - credible Explanations given to patients of treatment and control interventions: not stated Evidence of credibility: partial	Self assessed monthly symptom evaluation form to measure pain duration and intensity, onset of menses, amount and duration of flow, other symptoms (nausea, headache, backache, fluid retention, breast tenderness), and pain medication type and usage. "Monthly Pain Score" defined as sum of products of pain intensity and pain duration. Minimum important difference "arbitrarily defined" as average post-treatment score of less than half pre-treatment score Measures taken at baseline: monthly symptom evaluation form completed twice for past 6 months and most resent period (lowest monthly pain score was used as baseline) Measures taken at follow-up: monthly symptom evaluation form Evidence outcome measures are valid: partial – credible & clearly described but not validated Timing of follow up: monthly for 12 months Adverse events monitored: no	1. Improvement in pain score was found in 10/11 (90.9%) of the RA group, 4/11 (36.4%) of the PA group, 2/11 (22%) of the SC group, and 1/10 (10%) of the VC group. 2. Trends in medication usage: RA group 54% drop during, 41% drop post-treatment. PA group 24% drop during, 7% increase post-treatment. VC group 23% drop during, 7% increase post-treatment. SC group no change. 3. RA group "often reported improvement in nausea, headache, backache, fluid retention, breast tenderness". Do findings support conclusions: no - risk of bias and imprecision from small groups not noted by author	1. Risk of false positive results due to small sample size. 2 groups of 20 would have given more reliable results 2. No information about changes to menstrual regularity and bleed although aspect of monthly questionnaire. Limited information about other symptoms (headache, breast tenderness etc.)

Study	Design and Methods	Participants	Interventions	Outcomes	Findings	Notes
Ma 1999 Post-medication amenorrhoea #2	CT ***Study QA score: 23% (7.5/32)*** Justification for sample size: no Randomisation secure: none Recruitment procedures: not stated Recruitment time-period: not stated Number of participants: 215 Number in each group: Acupuncture group (AG) 117 Hormone treatment group (HTG) 98 Allocation concealment adequate: none Blinding: not stated Assessor blinding: not stated Baseline comparability: no Lost to follow-up: Not stated *(Table II suggests loss of 17: 7/117 from acupuncture group, 10/98 from control group?)* Intention to treat: by default if attrition 0% Appropriate statistical analyses: partial – actual probability values not given, few absolute numbers Alternative analyses possible: no	Location: China Source population: not stated Inclusion criteria: amenorrhoea with history of medication usage (COCP, Domperidone and Chlorpromazine), gynaecological exams to rule out organic disease. Exclusion criteria: not stated Characteristics of participants at intervention commencement: *Age*: AG age range 25-36 years, HTG 26-34 years *Ethnicity*: not stated *Socio-economic group*: not stated *Contraception*: not stated *Parity*: not stated *Obesity*: not stated *Severity*: not stated *Chronicity*: given as duration of symptoms: AG 0.8 to 7years, HTG 0.6 to 4years Number in subgroups: Stagnation of Liver Qi 134 Deficiency of Spleen and Kidney 31 Disharmony between Liver and Stomach 50 Distribution of principle confounders in subgroups: not stated	Electro-acupuncture formulae according to syndrome patterns compared to hormone treatment ***Interventions QA score: 41% (18/44)*** Setting: not stated Practitioner details: not stated Style of acupuncture: implicit Rationale for treatment: syndrome patterns Points used: REN4, REN3, Zigong, KID12, SP6, paravertebral points T3-L2 Plus relevant adjuvant points: *Stagnation of Liver Qi*: BL18, BL19, GB34 *Deficiency of Spleen and Kidney*: BL20, BL23, ST36 *Disharmony between Liver and Stomach*: BL18, BL20, ST36 Point locations: implicit Needling details: type and depth not stated Needle stimulation and response: Reducing manipulation for Excess conditions, reinforcing manipulation and warming needling for Deficiency. Deqi obtained then electricity applied to the needles. Number of treatments: 6 courses of 20 Needle retention time: 20 minutes Frequency of treatments: 5 days between courses Co-interventions: none Control intervention: *Hormone treatments:* Diethyl stilbestrol, Progesterone or Provera and Clomifene (for patients who wished to become pregnant) Treatment regimen: Diethyl stilbestrol 0.5mg for 20 days Progesterone 20mg qd or provera 10mg qd from day 16 for 5 days for 3-6 courses Clomifene 50mg - 250mg for 5-10 days each cycle, dependant upon response for 6 courses Rationale/sources that justify control: not stated Explanations given to patients of treatment and control interventions: not stated Evidence of credibility: not stated / credible	Physician assessed effective rate using: (1) Clinical records (2) Gynaecological exam (3) Ultra-sonographic exam *(4)* Radioimmunological study of PRL, FSH and LH hormone values *(see notes)* Effective rate is sum of A to C below: A. *Cure*: normal periods for 6 months, normal exam and test results, or pregnancy. B. *Marked effect*: slight menstrual irregularities (e.g. scanty bleed for 1-3 days) nearly normal exam and test results. C. *Effect*: menstrual irregularities (scanty bleed, breast tenderness, dysmenorrhoea), improvement in exam and test results. D. *No effect*: no change after 6 months of treatment. Measures taken at baseline: Clinical records Gynaecological exam Ultra-sonographic exam PRL, FSH and LH values Measures taken at follow-up: as above Evidence outcome measures are accurate: no Timing of follow-up: clinical symptoms recorded regularly, exams and tests at 1 month and 6 months Adverse events monitored: no	1. *Effective rate*: AG 94.9% (111/117) HTG 79.6% (78/98) with $p < 0.01$ at 6-month follow-up. 2. *Cured rate*: AG 61.8% (72/117) CG 43.9% (43/98) 3. *Hormone Values:* PRL, FSH and LH normalised in both groups, with values maintained closer to normal levels in the acupuncture group (normal values not given) Subgroup analyses: no Do findings support conclusions: no - risk of bias and imprecision from inadequate randomisation, distribution of prognostic variables, lack of validated outcome measures and observer bias not noted by author	*PRL = prolactin* *FSH = follicle stimulating hormone* *LH = luteinising hormone* 1. Frequency of treatments: unclear if daily or weekly 2. Positive element is attempt to validate practitioner assessed outcome with more credible measure of hormone values

Study	Design and Methods	Participants	Interventions	Outcomes	Findings	Notes
Thomas 1995 Primary dysmenorrhoea #4	CT Crossover design ***Study QA score: 55% (17.5/32)*** Justification for sample size: not stated Randomisation secure: none Recruitment procedures: not stated Recruitment time-period: not stated Number of participants: 31 Number in each group: Acupuncture group (AG) 19 TENS group (TG) 12 Allocation concealment adequate: no Blinding: limited information Assessor blinding: partial Baseline comparability: partial – details not provided *(see notes)* Lost to follow-up: 2 from acupuncture group ("despite describing benefit") Intention to treat: partial Appropriate statistical analyses: partial – within group change assessed but comparative change between groups was not, actual probability values not given, few absolute numbers Alternative analyses possible: partial – lack of simple outcome data	Location: Sweden Source population: patients attending Dept of Obst and Gynae at Karolinska Hospital and a private clinic Inclusion criteria: symptom duration > 5 years, previous ineffective treatment with NSAIDs or contraindicated due to aspirin allergy, gynaecological exam to rule out pathology, naïve to TENS or acupuncture Exclusion criteria: not stated Characteristics of participants at intervention commencement: *Age*: AG mean 30.4 years (SD 7.3), TG mean 27.8 years (SD 5.5) $p<0.05$ *Ethnicity*: not stated *Socio-economic group*: not stated *Contraception*: not stated *Parity*: AG 8/19, TG 9/12 *Obesity*: not stated *Severity*: not stated *Chronicity*: duration > 5 years	4 modes of acupuncture compared to 3 modes of TENS ***Interventions QA score: 79% (33/44)*** Setting: Dept of Obst and Gynae, Karolinska Hospital and a private clinic Practitioner details: long experience stated Style of acupuncture: not stated Rationale for treatment: spinal segments Points used: REN4, SP9, SP6, BL32 and 4 sites proximal to standard acu-points Point locations: implicit Needling details: needle type not stated, depth 1 to 3cms (morphology of site) Needle stimulation and response: *Mode 1*: Manual, deqi obtained every 5mins *Mode 2*: Low frequency electrical stimulation (2Hz) at 2 pairs of needles to give comfortable muscle contractions *Mode 3*: High frequency electrical stimulation (100Hz) at 2 pairs of needles to give comfortable parathesia *Mode 4*: Periosteal stimulation at tendinous insertions into bone, 30scs stimulation, pain elicited, stimulation repeated 3 to 4 times TENS treatments: *Mode 1*: Low frequency TENS 2Hz (LFT) *Mode 2*: High frequency TENS 100Hz (HFT) *Mode 3*: Placebo TENS (PT) Locations: T10 to L1 Pulse duration: 0.2ms Number of treatments: 2 treatments with each mode of acupuncture or TENS and preferred mode (10 for AG, 8 for TG) Treatment time: 20 minutes Frequency of treatments: 7 and 3 days prior to onset of menstruation for 5 (AG) or 4 (TG) consecutive months Rationale/sources that justify controls: to equitably distribute placebo between groups Explanations given to patients of treatment and control interventions: patients told placebo TENS involved an ultra-high frequency that might not be detected Evidence of credibility: partial – *(see notes)*	Self assessed questionnaire measuring menstrual pain (VAS), blood loss (scanty, moderate, heavy), nausea, vomiting, hours of work lost, analgesics taken and a subjective assessment of change (better, worse or no change) Measures taken at baseline: self assessed questionnaire for month prior to treatment Measures taken at follow-up: self assessed questionnaire Evidence outcome measures are valid: partial – credible and clearly described but not validated Timing of follow up: monthly post each mode of treatment (AG=5 months, TG=4 months), and 3-months post treatments or 8 and 7 months from baseline Adverse events monitored: no	1. AG: improvement in all measures, significant for pain, tablet intake and subjective assess, benefits maintained at follow-up. 2. TG: improvement in pain, tablet intake and subjective assess, with subjective assess maintained at follow up. Subgroup analyses: 1. LFT reduced pain and tablet use 2. HFT and PT no benefit 3. AG improvement continues over time despite sequence of modes differing amongst patients – evidence order not a confounding factor and results relate to efficacy of mode 4. AG no mode preferred TG 9/12 prefer LFT Do findings support conclusions: No – risk of bias and imprecision from inadequate randomisation, carry-over effects, risk of random error in sub-group analyses and observer bias not noted by authors.	*Evidence of control credibility: Inert TENS is a valid placebo. Credibility of periosteal stimulation as sham acupuncture on target population has not been demonstrated (see interventions)* 1. Distribution of the prognostic variables menarche, pain and "other treatment parameters" not provided 2. Acu-points that received electrical stimulation not clear – highlighted 6 but used 4. 3. Evidence of confounding from carry-over effects as overall improvement in the AG was cumulative. 5. 1 acupuncture and 1 TENS group would have given more valid results given sample size 4. Therapeutic effects from, and possible mechanisms for, periosteal stimulation noted.

Study	Design and Methods	Participants	Interventions	Outcomes	Findings	Notes
Ge 2000 Primary dysmenorrhoea #5	Case-series ***Study QA score: 41% (11.5/28)*** Recruitment procedures: not stated Recruitment time-period: not stated Number of participants: 108 Homogeneity addressed: yes Assessor blinding: not stated Lost to follow-up: not stated Follow-up long enough: partially Appropriate statistical analyses: Partial – explicitly reported by imperfect X^2 test, and one implicit but correct t test Other evidence of quality assurance: no	Location: China Source population: female students in a teacher training school Inclusion criteria: implicit - female students in teacher training school with dysmenorrhoea Exclusion criteria: not stated Characteristics of participants at intervention commencement: *Age*: 17 to 19 years *Ethnicity*: not stated *Socio-economic group*: not stated *Contraception*: not stated *Parity*: *Obesity*: not stated *Severity*: 40 severe pain, 42 moderate, 21 mild pain *Chronicity*: 0.5 to 4.0 years, mean 3.2. *Pain onset*: 39 experienced pain prior to, 45 during, 24 post-menstruation *TCM diagnosis*: Excess syndrome 81 Deficient syndrome 27 Number in sub-groups: 37/108 participants 10 normal controls Distribution of principle confounders in sub-groups: *Severity*: participants had mild to moderate dysmenorrhoea	Acupuncture formula and cupping according to TCM identification ***Interventions QA score:*** Setting: implicit - No. 463 Hospital, Liaoning Province Practitioner details: not stated Style of acupuncture: implicit Rationale for treatment: TCM identification Points used: BL26, SP6, REN4 Point locations: not stated Needling details: BL26: reusable needle gauge 26, length 2 cun, depth 1.5cun. SP6 & REN4: dependent upon TCM identification (no details) Needle stimulation and response: BL26: reducing method, deqi obtained. SP6 & REN4: dependent upon TCM identification (no details) Number of treatments: not stated Needle retention time: 30 minutes Frequency of treatments: daily for the week prior to 3 menstrual cycles Co-interventions: Cupping for 10mins at BL26, once needle withdrawn	Physician assessed effective rate using: (1) Clinical observation (2) PGF2 values in subgroup *(see notes)* The effective rate is the sum of items A to C: *A. Cured*: symptoms disappeared with no relapse in 3-months. *B. Markedly effective:* symptoms basically disappeared *C. Improved*: symptoms were obviously improved *D. Failed*: no obvious improvement. Measures taken at baseline: PGF2 values in subgroup Measures taken at follow up: PGF2 values in subgroup Evidence outcome measures are accurate: partial – states that used method from the Beijing CooP Group, kit provided by Animal Institute of the Chinese Academy of Sciences, no evidence as to validity for this patient group Timing of follow up: 3 months Adverse events monitored: no	1. Physician assessed total effective rate: 91.6% (99/108) with 55.5% (60/108) cured 2. Severity of symptoms did not affect outcome 3. Patients with excess syndromes responded better than those with a deficiency syndrome Sub-group findings: 1. Prior to treatment, PGF2 levels higher than those of 10 normal controls ($p>0.01$) 2. PGF2 levels lowered and approached normal levels post treatment ($p>0.05$). Do findings support conclusions: no – risk of bias and imprecision inherent in study design not noted. Risk of bias from lack of validated outcomes and observer bias not noted with causal hypotheses made.	*PGF2 values: test to determine prostaglandin F2 levels in menstrual blood* *(see outcomes)* 1. No rationale for excluding 71/108 (66%) of sample from PGF2 tests 2. PGF2 positive addition as way of validating practitioner assessed outcomes

Study	Design and Methods	Participants	Interventions	Outcomes	Findings	Notes
Yu 1998 Primary dysmenorrhoea #6	Case-series ***Study QA score:8% (2/24)*** Recruitment procedures: not stated Recruitment time-period: not stated Number of participants: 32 Homogeneity addressed: no Assessor blinding: no Lost to follow-up: not stated Follow-up long enough: no Appropriate statistical analyses: NA Alternative analyses possible: NA Other evidence of quality assurance: no :	Location: China Source population: outpatients from acupuncture and gynaecology departments Inclusion criteria: not stated Exclusion criteria: genital organ disease discovered with gynaecological exam Characteristics of participants at intervention commencement: *Age*: 17 to 29 years, mean 21.3 years *Ethnicity*: not stated *Socio-economic group*: not stated *Contraception*: not stated *Parity:* not sated *Obesity*: not stated *Severity*: not stated *Chronicity*: symptom duration of 2 days to 4 years	Acupuncture formulae ***Interventions QA score:29% (10/34)*** Setting: implicit - Zhujiang Hospital, First Medical Military University, Guangdong Province Practitioner details: not stated Style of acupuncture: not stated Rationale for treatment: not stated Points used: REN3, SP10, SP6 and if pain severe BL32 Point locations: not stated Needling details: needle tip 45% angle towards perineum, needle type and depth of insertion not stated Needle stimulation and response: REN3: strong deqi sensation in lower abdomen, manipulated every 5min for 25 min. SP10 & SP6: reducing method. Number of treatments: not stated Needle retention time: 25 minutes Frequency of treatments: Up to 3 daily sessions prior to menstruation *(see notes)* Co-interventions: Electo-acupuncture if pain severe – no further details provided	Physician assessed effective rate as sum of items A & B: *A. Remarkably effective:* pain disappears completely, patient resumes normal work and life *B. Effective*: pain is lightened and patient may resume normal work *C. No effect*: no obvious relief. Measures taken at baseline: not stated Measures taken at follow-up: not stated Evidence outcome measures are accurate: no Timing of follow-up: not stated Adverse events monitored: yes – Treatment may not be acceptable to many women due to strong deqi sensation required	1. Physician assessed total effective rate 93.8% (30/32), 2. Remarkably effective rate 65.1% (21/32) 3. 1 treatment effective in 25 cases and 2 treatments in 5 cases. 4. No effect in the 2 patients unwilling to experience strong deqi sensation. Do findings support conclusions: no – risk of bias and imprecision inherent in study design not noted. Risk of bias from lack of validated outcomes and observer bias not noted. Causal hypotheses made.	1. Precise treatment regimen unclear. Case report was of a patient advised to return for monthly acupuncture treatments 2-3 days prior to every menstruation.

Study	Design and Methods	Participants	Interventions	Outcomes	Findings	Notes
Jiao 1997 Primary dysmenorrhoea #7	Case-series ***Study QA score:2% (0.5/24)*** Recruitment procedures: not stated Recruitment time-period: not stated Number of participants: 30 Homogeneity addressed: no Assessor blinding: not stated Lost to follow-up: not stated Follow-up long enough: 0 Appropriate statistical analyses: NA Clarity of findings: no – merged criteria and results Other evidence of quality assurance: no	Location: China Source population: not stated Inclusion criteria: not stated Exclusion criteria: not stated Characteristics of participants at intervention commencement: *Age*: 15 to 27 years *Ethnicity*: not stated *Socio-economic group*: not stated *Contraception*: not stated *Parity*: not sated *Obesity*: not stated *Severity*: not stated *Chronicity*: symptom duration from 2 months to 7 years *Marital status*: 25 single, 5 married	Acupuncture formulae and lifestyle advice according to syndrome patterns ***Interventions QA score:44% (15/34)*** Setting: implicit – Central Hospital of Miaoshan, Shandong Province Practitioner details: not stated Style of acupuncture: not stated Rationale for treatment: implicit – syndrome patterns Points used: REN3, SP6, BL32 with appropriate adjuvant points: *Qi Stagnation and Blood Stasis* - SP1, LI4 *Cold-Damp in Congelation* - DU4, REN4 *Deficiency of Qi and Blood* - REN6, ST3 *Insufficiency of Liver and Kidney* - BL23, BL18 Point locations: not stated Needling details: needle type and depth of insertion not stated Needle stimulation and response: Moderate needle stimulation, manipulated at intervals to send sensation to abdomen and perineum Number of treatments: not stated Needle retention time: 30 minutes Frequency of treatments: 10 daily from 7 days prior to menstruation Co-interventions: Lifestyle advice: avoid raw, cold food during menstruation, be in good spirits and personal health	Physician assessed effective rate as sum of items A to C: *A. Cured*: pain entirely relived *B. Markedly effective*: pain basically relieved *C. Effective*: pain less Measures taken at baseline: not stated Measures taken at follow-up: not stated Evidence outcome measures are accurate: no Timing of follow-up: not stated Adverse events monitored: no	1. Physician assessed total effective rate 100% *(30/30)* 2. Cure rate 86.6% (26/30) Do findings support conclusions: partial – No causal hypotheses made, but risk of bias and imprecision inherent in study design not noted. Risk of bias from lack of validated outcomes and observer bias not noted.	

Study	Design and Methods	Participants	Interventions	Outcomes	Findings	Notes
Wang 1996 Primary and secondary dysmenorrhoea #8	Case-series ***Study QA score:27% (6.5/24)*** Recruitment procedures: not stated Recruitment time-period: not stated Number of participants: 30 Homogeneity addressed: no Assessor blinding: not stated Lost to follow-up: not stated Follow-up long enough: yes Appropriate statistical analyses: NA Alternative analyses possible: NA Other evidence of quality assurance: no	Location: China Source population: Outpatients from Fugou County Hospital Inclusion criteria: implicit Exclusion criteria: not stated Characteristics of participants at intervention commencement: *Age*: 17-52 years, 19 = 17-25 years 25 = 26-30 years 28 = 31-35 years 35 = 36-40 years 15 = 41-52 years *Ethnicity*: not stated *Socio-economic group*: not stated *Contraception*: not stated *Parity:* not sated *Obesity*: not stated *Severity*: not stated *Chronicity*: symptom duration from 2 to 35 years *Diagnosis*: 64/122 primary dysmenorrhoea, 58/122 secondary dysmenorrhoea (criteria not stated). *TCM diagnosis*: 25/122 deficiency type 38/122 excess type 35/122 deficiency-cold type 24 /122 excess-heat type.	Acupuncture formula, patient administered moxabustion and lifestyle advice ***Interventions QA score:60% (21.5/36)*** Setting: Fugou County Hospital, Henan Province Practitioner details: not stated Style of acupuncture: Traditional Chinese Medicine (TCM) Rationale for treatment: TCM theory, channel theory and point actions described Points used: REN3 and SP6 Point locations: implicit Needling details: REN3: 2-cun oblique insertion, SP6: 2.5-cun insertion. Needle type not stated Needle stimulation and response: REN3: manipulated every 5 min for 30 min, sensation to perineal region SP6: moderately strong reinforcing-reducing stimulation, manipulated every 5 min for 30 min, sensation to lower leg and abdomen Number of treatments: 10 to 14 Needle retention time: 30 minutes Frequency of treatments: 2 courses of 5-7 daily treatments, 2 days between each course Co-interventions: (1) Patient administered moxabustion: moxa post acupuncture held for 20min 2-cun above REN3 until skin red and congested (2) Lifestyle advice: 1 week prior to menstruation pay attention to hygiene, avoid cold food, keep warm, refrain from sexual intercourse (N.B. case report involved cupping)	Physician assessed effective rate as sum of items A & B: *A. Curative*: all signs and symptoms disappeared, no recurrence after 1 year *B. Effective*: remarkable improvement in signs and symptoms *C. Ineffective:* no change Measures taken at baseline: not stated Measures taken at follow-up: not stated Evidence outcome measures are accurate: no Timing of follow-up: 1 year Adverse events monitored: no	1. Physician assessed total effective rate 95.9% (117/122) 2. Curative rate 80.3% (98/122) 3. The earlier the treatment the better the results Do findings support conclusions: no – definitive causal hypotheses made, risk of bias and imprecision inherent in study design not noted. Risk of bias from lack of validated outcomes and observer bias not noted.	1. Title of study is misleading as "needling and cupping …"

Study	Design and Methods	Participants	Interventions	Outcomes	Findings	Notes
Yong-hai 1995 Primary dysmenorrhoea #9	Case-series ***Study QA score: 2% (0.5/24)*** Recruitment procedures: not stated Recruitment time-period: over 20 years Number of participants: 30 Homogeneity addressed: no Assessor blinding: not stated Lost to follow-up: not stated Long enough follow-up: no Appropriate statistical analyses: NA Clarity of findings: no – did not define "cure" Other evidence of quality assurance: no	Location: China Source population: not stated Inclusion criteria: not stated Exclusion criteria: not stated Characteristics of participants at intervention commencement: *Age*: 16 to 20 years *Ethnicity*: not stated *Socio-economic group*: not stated *Contraception*: not stated *Parity*: not sated *Obesity*: not stated *Severity*: not stated *Chronicity*: symptom duration from 1 to 5 years 20/30 1 year 6/30 2-3 years 4/30 4-5 years *Marital status*: unmarried	Acupuncture formula ***Interventions QA score:36% (11.5/32)*** Setting: implicit - Clinic of the Second Geological Team, Hebei Province Practitioner details: not stated Style of acupuncture: not stated Rationale for treatment: TCM diagnosis Points used: REN3, 4, 6, SP6, ST36 Point locations: implicit Needling details: needle type and depth not stated Needle stimulation and response: *Moderate symptoms*: mild reinforcing and reducing method *Excess syndrome, strong constitution, obstinate pain*: reducing method *Insufficiency syndrome, weak constitution*: reinforcing method. Number of treatments: not stated Needle retention time: 30 minutes Frequency of treatments: 15 weekly per course, up to 5 courses	Physician assessed effective rate as sum of items A to C: A. Cure after 1 course of treatment B. Cure after 3 courses of treatment C. Relief after 5 courses of treatment Measures taken at baseline: not stated Measures taken at follow-up: not stated Evidence outcome measures are accurate: no Timing of follow-up: not stated Adverse events monitored: no	1. Physician assessed total effective rate 100% (30/30) 2. Cure rate 60% (18/30) Do findings support conclusions: partial – no definitive causal hypotheses made, but risk of bias and imprecision inherent in study design, from lack of validated outcomes and observer bias not noted	Did not even define a "cure"

Study	Design and Methods	Participants	Interventions	Outcomes	Findings	Notes
Tureanu 1994 Post-oral contraceptive menstrual irregularities and amenorrhoea #10	Case-series ***Study QA score: 33% (8/24)*** Recruitment procedures: not stated Recruitment time-period: not stated Number of participants: 38 Homogeneity addressed: no Assessor blinding: not stated Lost to follow-up: not stated Long enough follow-up: partial Appropriate statistical analyses: NA Alternative analyses possible: NA Other evidence of quality assurance: no	Location: Romania Source population: outpatient clinic Inclusion criteria: menstrual disorders and history of oral contraceptive (OC) use within past 5 years Exclusion criteria: Chronic and metabolic diseases or post-traumatic disorders (cervical stenosis, uterine adhesions etc.) Characteristics of participants at intervention commencement: *Age*: 24 to 40 years 11/38 age 24-30 11/38 age 31-35 16/38 age 36-40 *Ethnicity*: not stated *Socio-economic group*: not stated *Contraception*: not stated *Parity*: not sated *Obesity*: not stated *Severity*: not stated *Chronicity*: 3-8 irregular or missed cycles, mean 5 20/30 1 year 6/30 2-3 years 4/30 4-5 years *Length of OC usage*: 1.6-5 years, mean 3 *Length of menstrual cycle*: 9/38 32-45 days 5/38 21-26 days, 2/38 irregular 12/38 amenorrhoea *Menstrual problems prior to OC usage*: 20/38 14/20 oligomenorrhea 6/20 irregular cycles	Individualised TCM acupuncture ***Interventions QA score: 55% (17.5/32)*** Setting: not stated Practitioner details: not stated Style of acupuncture: TCM Rationale for treatment: syndrome patterns Typical points: *Irregular Menstruation:* (1) Heat in the Blood: REN3, LI11, SP10, KID5 or REN3, LI43, SP10, KID2, LIV2 (2) Qi Deficiency: REN6, SP6, ST36, REN12 or REN6, SP4, KID3, ST36 (3) Blood Deficiency: REN4, SP6, BL17, BL20 or REN6, ST36, SP6, BL17 (4) Cold in the Blood: REN6, SP6, REN4, ST29 Qi Stagnation: REN6, ST25, SP8, KID13 or REN6, ST29, LIV3, P6 (5) Kidney Deficiency: REN4, BL23, SP6, ST36, KID8 or KID10 (6) Liver Qi Stagnation: REN6, LIV5, KID14, P5 or SP6, LIV3, REN6, P6 *Amenorrhoea:* (1) Blood Stagnation: REN3, SP10, LIV3, SP6, ST29, LI4 (2) Blood Deficiency: REN4, SP6, ST36, BL20, BL18 or REN6, SP4, BL23, LI4 Supplementary points: ST30, LIV14, REN17, HT7, LIV2, KID10 to treat individual symptoms (palpitations etc.) Point locations: implicit - classical Needling details: Non-disposable silver filiform: 25,40,65mm; gauge 0.30,32,40 Needle stimulation and response: 1) Heat in the Blood, Qi Stagnation, Blood Stagnation: Reducing method (2) Qi Deficiency, Blood Deficiency, Kidney Deficiency: Reinforcing method (3) Cold in the Blood, Liver Qi Stagnation: Even method. Response not stated. Number of treatments: not stated Needle retention time: 15 to 20 minutes Frequency of treatments: Daily 4-5 days prior to predicted onset of menstruation and 4-5 days post per course, up to 3 courses.	Physician assessment according to: (1) General condition: presence, regularity and character of menstrual flow (2) Basal body temperature (BBT) (3) Mid-cycle levels of luteinizing hormone (Discretest) (4) Insler score Measures taken at baseline: not stated Measures taken at follow-up: not stated Evidence outcome measures are accurate: partial – BBT and mid-cycle LH are recognisable clinical outcomes but no evidence of validity for this patient group Timing of follow-up: average of 5 cycles Adverse events monitored: no	1. Regular menstrual cycles resumed in 84.2% (32/38), 15 after one course of treatment, 9 after 2, 8 after 3. 2. Ovulation regulated as evidenced by mid-cycle LH values and cervical mucus studies 3. Patients with history of menstrual irregularities and those who'd used OCP for more than 2-3 years needed more treatment Do findings support conclusions: no – authors state results evidence of efficacy. Risk of bias and imprecision inherent in study design not noted, lack of validated outcomes, absence of defined data collection and analysis methods, and risk of observer bias not noted.	Positive feature of more credible outcomes in an attempt to validate practitioner assessments

Study	Design and Methods	Participants	Interventions	Outcomes	Findings	Notes
Zhang 1994 Dysfunctional uterine bleeding (heavy or prolonged bleed with haemorrhagic anaemia) #11	Case-series ***Study QA score:13% (3/24)*** Recruitment procedures: not stated Recruitment time-period: not stated Number of participants: 50 Homogeneity addressed: no Assessor blinding: not stated Lost to follow-up: not stated Follow-up long enough: partially Appropriate statistical analyses: NA Alternative analyses possible: NA Other evidence of quality assurance: no	Location: China Source population: not stated Inclusion criteria: Adolescent girls and women of childbearing age with heavy or prolonged menstrual bleeding accompanied by haemorrhagic anaemia Exclusion criteria: Menopausal women, organic changes, diseases of whole body, iatrogenic factors Diagnosis determined by measurement of hormones, BBT and endometrial biopsy Characteristics of participants at intervention commencement: *Age:* 17 to 37 years *Ethnicity:* not stated *Socio-economic group:* not stated *Contraception:* not stated *Parity:* not stated *Obesity:* not stated *Severity:* not stated *Chronicity:* symptom duration 6 months to 20 years	Acupuncture formula ***Interventions QA score: 39% (12.5/32)*** Setting: not stated Practitioner details: not stated Style of acupuncture: not stated Rationale for treatments: TCM diagnosis Points used: Points equate to LI4, BL40, BL29 and DU20 Point locations: zones, lines and anatomy Needling details: needle type and depth of insertion not stated Needle stimulation: (1) Qi deficiency type: tonification and reinforcement method (2) Blood heat type: heat-clearing method (3) Blood stasis type: reduction and purgation method Needle response: not stated Number or treatments: not stated Needle retention time: not stated Frequency of treatments: 10 treatments every other day	Physician assessed effective rate as sum of items A to C: *A. Cured*: bleeding stops, normal menstruation, correction of anaemia *B. Markedly effective*: marked improvement of symptoms *C. Effective*: improvement of symptoms *D. Ineffective:* no improvement Measures taken at baseline: not stated Measures taken at follow-up: not stated Evidence outcome measures are accurate: no Timing of follow-up: 6-months Adverse events monitored: no	1. Physician assessed effective rate 100% (50/50) 2. Cured rate 86% (43/50) Do findings support conclusions: partial – no definitive causal hypotheses made, but risk of bias and imprecision inherent in study design, from lack of validated outcomes and observer bias not noted.	

Study	Design and Methods	Participants	Interventions	Outcomes	Findings	Notes
Mo 1993 Ovulatory dysfunction #19	Case-series ***Study QA Score:36% (10/28)*** Recruitment procedures: not stated Recruitment time period: 12 months Number of participants: 34 Homogeneity addressed: no Assessor blinding: not stated Lost to follow-up: not stated Follow-up long enough: no Intention to treat: no Appropriate statistical analyses: partial - simple summary data provided but subgroup to small to test for significance and actual probability values not given Clearly described: 3 clear tables Other evidence of quality control: no	Location: China Source population: not stated Inclusion criteria: 1. Secondary one-degree amenorrhoea (responds to progesterone) or anovulatory DUB with infertility due to anovulation 2. Monophasic BBT for previous 3 months 3. No follicular development or anovulation as evidenced by B ultrasonic exam between days 12 and 14 Exclusion criteria: not stated Characteristics of participants at intervention commencement: *Age*: 21-39 years, mean 30 *Ethnicity*: not stated *Socio-economic group*: not stated *Contraception*: not stated *Parity*: not stated *Obesity*: not stated *Severity*: not stated *Chronicity*: symptom duration 1 to 12 years, mean 6.5 years *Marital status*: 20/34 single, 14/34 married Number in subgroups: 10/34 participants 7 normal controls Distribution of principle confounders in subgroups: not stated	Acupuncture formulae ***Interventions QA Score: 45% (1.5/32)*** Setting: out-patient department of Zhejiang College of TCM, Hangzhou Practitioner details: not stated Style of acupuncture: TCM Rationale for treatments: TCM theory Points used: 2 formulae used alternately: (1) REN4, Zigong, SP6 (2) BL18, BL23, Shiqixhui, SP6 Needling details: type of needle and depth of insertion not stated Needle stimulation and response: Uniform reducing-reinforcing until deqi obtained Number of treatments: not stated Needle retention time: 20 ~ 30 minutes Frequency of treatments: 3 weekly for 3 months constituted 1 course of therapy	Physician assessed effective rate using: (a) Clinical notes (b) BBT (c) B ultrasonic exam (scan) (d) FSH, LH, E2, P, PP levels *(see notes)* The effective rate is the sum of A & B below: A. Markedly effective: menstruation with normal menstrual volume for 2 cycles or pregnancy as evidenced by scan and biphasic BBT B. Effective: monophasic BBT but with evidence of change as evidenced by blood loss and follicular development C. Ineffective: no change Measures taken at baseline: BBT and scan (FSH, LH, E2, P, PP levels in subgroup) Measures taken at follow-up: BBT (scan and FSH, LH, E2, P, PP levels in sub-group) Timing of follow-up: not stated Evidence outcome measures are accurate: partial – tests are valid, hormone levels tested according to WHO guides. Adverse events monitored: no	1. Total effective rate of 82% (28/34) 2. Markedly effective rate 35% (12/34), effective 48% (16/34), ineffective 18% (6/34) 5. 2 pregnancies 6. Excessive menstrual loss normalised 7. At baseline, 2/34 had atypically biphasic BBT, and 32 monophasic BBT 8. Post treatment 7/34 had typical biphasic BBT, 7/34 atypical biphasic, and 20/34 monophasic 9. At baseline scan revealed no follicular development Subgroup findings: 10. Scan post treatment revealed 15/20 had follicular development, 8 with maturity comparable to controls ($p>0.01$) 11. Regulation of hormones observed ($p>0.05$) Do findings support conclusions: no – results presented as evidence of effectiveness. Risk of bias and imprecision inherent in study design, from lack of validated outcomes, random error and observer bias not noted	*FSH= follicle stimulating hormone* *LH= luteinizing hormone* *E2= estradiol* *P= progesterone* *PP= pituitary prolactin* *(see outcomes)* 1. Inclusion of 2 patients with DUB confuses the definition of the study population 2. Evidence scan between days 12 to 14 unreliable as no follicular development revealed at baseline despite biphasic BBT in 7 participants 3. No rationale for excluding 20/34 from blood endocrine analyses and follow-up B ultrasonic exams 4. But positive addition as attempt to validate practitioner assessments

Study	Design and Methods	Participants	Interventions	Outcomes	Findings	Notes
Liu 1992 Amenorrhoea 12#	Case-series ***Study QA Score: 25% (6/24)*** Recruitment procedures: not stated Recruitment time period: 8.5 years Number of participants: 225 Homogeneity addressed: yes Assessor blinding: not stated Lost to follow-up: not stated Follow-up long enough: partially Appropriate statistical analyses: NA Clarity of findings: no, draws statistical inferences without necessary tests Other evidence of quality control: no	Location: China Source population: not stated Inclusion criteria: not stated Exclusion criteria: not stated Characteristics of participants at intervention commencement: *Age*: 19-39 years *Ethnicity*: not stated *Socio-economic group*: not stated *Contraception*: not stated *Parity*: not stated *Obesity*: not stated *Severity*: not stated *Chronicity*: symptom duration 3 months to 1.4 years *Marital status*: 39/225 single, 186/225 married *TCM diagnosis*: Liver & Kidney Deficiency 6 Deficiency of Qi & Blood 63 Stagnancy of Qi & Blood Stasis 135 Phlegm Stagnancy 21	Acupuncture formulae and moxabustion according to TCM diagnosis ***Interventions QA Score: 37% (12.5/34)*** Setting: implicit – Shangqiu County Hospital, Henan Province Practitioner details: not stated Style of acupuncture: TCM Rationale for treatments: syndrome patterns Points used: Liver & Kidney Deficiency: BL18, BL23, SP6, REN6, DU4 Deficiency of Qi and Blood: SP10, ST36 Stagnancy of Qi and Blood Stasis: SP6, REN6, SP8, P6, LIV3, SJ6 Phlegm Stagnancy: REN3, SP6, REN6, SP9, ST40, SP10 Needling details: type of needle and depth of insertion not stated Needle stimulation: Liver & Kidney Deficiency, Deficiency of Qi and Blood, and Stagnancy of Qi and Blood Stasis: reinforcing method. Needles manipulated 3-5 times during treatment. Phlegm Stagnancy: not stated Needle response: not stated Number of treatments: unclear, case report of 10 daily treatments Needle retention time: 20 – 30 minutes Frequency of treatments: 5 daily treatments constituted 1 course of treatment Co-interventions: Deficiency of Qi & Blood: moxabustion at SP6, REN6, BL20 (no further details)	Physician assessed effective rate as sum of items A & B: *A. Cured*: restoration of menstruation, disappearance of symptoms and signs for over 6 months. *B. Effective*: restoration of menstruation with small amount of blood, disappearance of main symptoms and signs. *C. Failed*: no fundamental change. Measures taken at baseline: not stated Measures taken at follow-up: Evidence outcome measures are accurate: no Timing of follow-up: not stated Adverse events monitored: no	1. Physician assessed total effective rate 98% (221/225) 2. Cure rate 72% (162/225) 3. 99% with a symptom duration of less than 6 months responded to treatment (97/121) 4. 97% with a symptom duration of more than 6 months responded to treatment (65/104) 5. Effective rate according to syndrome patterns: Liver & Kidney Deficiency 83% (5/6) Deficiency of Qi and Blood 100% (63/63) Stagnancy of Qi & Blood Stasis 98% (133/135) Phlegm Stagnancy 95% (20/21) Do findings support conclusions: no – results presented as evidence of effectiveness. Risk of bias and imprecision inherent in study design; from lack of validated outcomes and observer bias not noted	

Study	Design and Methods	Participants	Interventions	Outcomes	Findings	Notes
Liang 1990 Primary dysmenorrhoea #13	Case-series ***Study QA Score:16% (4.5/26)*** Recruitment procedures: not stated Recruitment time period: not stated Number of participants: 106 Homogeneity addressed: yes Assessor blinding: not stated Lost to follow-up: not stated Follow-up long enough: partially Appropriate statistical analyses: NA Clarity of findings: no, draws statistical inferences without necessary tests Other evidence of quality control: no	Location: China Source population: not stated Inclusion criteria: not stated Exclusion criteria: not stated Characteristics of participants at intervention commencement: *Age*: 15 to 41 years *Ethnicity*: not stated *Socio-economic group*: not stated *Contraception*: not stated *Parity*: not stated *Obesity*: not stated *Severity*: not stated *Chronicity*: symptom duration from less than five to more than 6 years *Marital status*: 66/106 unmarried, 40/106 married *Length of menstrual cycle and bleed*: 20/106 irregular 50/106 cycle of 3-8/30 days, 36/106 cycle of 3-6/26-28 days.	Acupuncture formula with plum-blossom-like needle and lifestyle advice ***Interventions QA Score: 30%(9.5/32)*** Setting: implicit – Hospital of Peking Union Medical College Practitioner details: not stated Style of acupuncture: not stated Rationale for treatments: not stated Points used: LIV 2, SP 1, SP 4, LIV 3, SP 6, REN4 Needling details: Plum-blossom-type needle (no further details provided) Needle stimulation: needle tapped at 70 to 90 beats per minute with wrist force Needle response: not stated Number of treatments: 9 Duration of needle tapping at points: not stated Frequency of treatments: 3 daily treatments 3 days prior to menstruation for 3 cycles Co-interventions: Lifestyle advice: avoid uncooked and cold food and cold water during menses	Physician assessed effective rate as sum of items A to C: *A. Cured*: after 3 treatments no discomfort at menstruation for 3 cycles. *B. Marked effective*: after 3 treatments patient experiences slight menstrual discomfort followed by 2 pain free cycles. *C. Improved:* after 6 treatments no pain is experienced at 3rd cycle. *D. Ineffective*: after 9 treatments the pain has been relieved but not resolved Measures taken at baseline: not stated Measures taken at follow-up: not stated Evidence outcome measures are accurate: no Timing of follow-up: 3 months Adverse events monitored: no	1. Physician assessed total effective rate 89% (94/106), total cure 28% (30/106) 2. Total effective rate according to age: <29 years 93%, >30 years 79% 3. Total effective rate according to symptom duration: >6 years 81%, <5 years 92%. Do findings support conclusions: no – results presented as evidence of effectiveness. Risk of bias and imprecision inherent in study design, from lack of validated outcomes and observer bias not noted	

Study	Design and Methods	Participants	Interventions	Outcomes	Findings	Notes
Zhan 1990 Primary dysmenorrhoea #14	Case-series ***Study QA Score:25% (6/24)*** Recruitment procedures: not stated Recruitment time period: not stated Number of participants: 32 Homogeneity addressed: partial Assessor blinding: not stated Lost to follow-up: not stated Follow-up long enough: yes Appropriate statistical analyses: NA Clarity of findings: no, draws statistical inferences without necessary tests Other evidence of quality control: no	Location: China Source population: not stated Inclusion criteria: not stated Exclusion criteria: not stated Characteristics of participants at intervention commencement: *Age*: 15 to 42 years 20/32 20-30 years *Ethnicity*: not stated *Socio-economic group*: not stated *Contraception*: not stated *Parity*: not stated *Obesity*: not stated *Severity*: not stated *Chronicity*: symptom duration a few days to10 years *TCM diagnosis*: all were Excess type: 20/32 Cold and Damp Retention 12/32 Qi Stagnation and Blood Stasis	Acupuncture formula according to TCM diagnosis ***Interventions QA Score: 49% (16.5/34)*** Setting: implicit – Chenjiaqiao Hospital, Shanghai Practitioner details: not stated Style of acupuncture: TCM Rationale for treatments: syndrome patterns Points used: *Cold and Damp Retention:* LI4, SP6 *Qi Stagnation and Blood Stasis*: LI4, SP6, REN6 Needling details: needle type and depth of insertion not stated Needle stimulation: reducing method, needles manipulated every 3 to 5 minutes Needle response: not stated Number of treatments: not stated Needle retention time: 20 minutes Frequency of treatments: Up-to 6 daily treatments: 3 prior to menstruation followed by a) 3 at onset of menstruation or b) one week prior to next cycle Co-interventions: Moxabustion at REN4 for Cold and Damp Retention	Physician assessed effective rate as sum of items A & B: *A. Cured:* pain ceased after 1 treatment. Course of treatment given to consolidate effects *(see notes)*. No recurrence at 6 months follow-up. *B. Effective*: marked alleviation after 1 treatment, 1 to 3 courses for consolidation, mild pain at later menstruations. *C. Ineffective*: no alleviation of pain after 3 daily treatments. Measures taken at baseline: not stated Measures taken at follow-up: not stated Evidence outcome measures are accurate: no Timing of follow-up: 6-months Adverse events monitored: no	1. Physician assessed total effective rate 97% (31/32), total cure 63% (20/32) 2. Total effective rate according to TCM diagnosis: Cold and Damp Retention 95% (19/20), Qi Stagnation and Blood Stasis 100% (12/12) Do findings support conclusions: no – results presented as evidence of efficacy. Risk of bias and imprecision inherent in study design, from lack of validated outcomes and observer bias not noted	1. Unclear what constitutes a course of treatment. The 2 cases described received 3 and 9 sessions.

Study	Design and Methods	Participants	Interventions	Outcomes	Findings	Notes
Wangcheng 1988 Functional uterine bleeding #15	Case-series ***Study QA Score:32% (9/28)*** Recruitment procedures: not stated Recruitment time period: not stated Number of participants: 30 Homogeneity addressed: no Assessor blinding: not stated Lost to follow-up: 10 - 20/30 followed-up for 3 cycles, attrition not accounted for Follow-up long enough: no Appropriate statistical analyses: partial – t tests reported implicitly but correctly Clarity of findings: yes, two clear tables Other evidence of quality control: no	Location: China Source population: in-patients and out-patients Inclusion criteria: not stated Exclusion criteria: other possible causes of uterine bleeding e.g. blood or liver disease, hypertension, gynaecological inflammation, tumours, injuries or foreign bodies (including IUD's) Characteristics of participants at intervention commencement: *Age*: 18-52 years, mean 36 *Ethnicity*: not stated *Socio-economic group*: not stated *Contraception*: not stated *Parity*: not stated *Obesity*: not stated *Severity*: not stated *Chronicity*: symptom duration: mode 1-3 years, 6/30 < 6 months, 4/30 >10 yrs *Treatment:* no herbs or drug therapy for 3 days *Menstrual regularity*: 7/30 short cycle (<21 days) 8/30 prolonged (>35 days) 15/30 irregular *Menstrual flow*: 9-15 days in 12/30, 16-30 days in 13/30, >30 days 5/30 *Age at menarche*: 13-18 years, mode 15 *Infertility*: 4/30 *Artificial abortion*: 15/30 *Marital status*: 4/30 unmarried *Biopsy of endometrium*: in 16/30 showed 15/16 anovulatory, 1/16 ovulatory. BBT in 14/30 showed 11/14 anovulatory, 3/14 ovulatory	Acupuncture formula according to TCM diagnosis ***Interventions QA Score: 47% (15/32)*** Setting: not stated Practitioner details: not stated Style of acupuncture: not stated Rationale for treatments: syndrome patterns Points used: *Deficiency in Spleen*: REN 6, Ren 4, SP 6, BL 20, BL 23 *Stagnancy in Liver*: LIV 3, SP 6, REN 4, BL 18, BL 23. *Deficiency in Heart and Spleen*: HE 7, REN 6, REN 4, SP 6, BL 15, BL 20, BL 23. *Deficiency of Yin of Liver and Kidney*: REN 4, LIV 5, KID 12, BL 18, BL 23, DU 4. *To improve the general condition*: ST 36, BL 20, BL 23, DU 4 Needling details: needle type and depth of insertion not stated Needle stimulation: Deficiency in Spleen: reinforcing method. Stagnancy in Liver: Reducing method, manipulated every 10mins Deficiency in Heart and Spleen and Deficiency of Yin of Liver and Kidney: reinforcing method, manipulated every 10mins Needle response: deqi obtained Number of treatments: not stated Needle retention time: 30 minutes Frequency of treatments: 15 daily per course, 3 days between courses	Physician assessed effective rate using: (1) Clinical observation (2) Change of blood picture in subgroup 20/30 (3) Haematological indices in subgroup of 8/30 with anovulatory FUB compared to 7 normal controls Effective rate is sum of items A & B: *A. Marked effectiveness*: bleeding stopped, symptoms relieved with 7 treatments. Regular menstruation with normal bleed for 3 cycles. *B. Effectiveness*: bleeding stopped, symptoms relieved with 10 treatments. Regular menstruation with slightly heavy bleed for 2 cycles. *C. Failure*: No benefit after 10 treatments. Measures taken at baseline: Change of blood picture 20/30, Haematological indices 8/30 Measures taken at follow-up: as above Evidence outcome measures are accurate: no Timing of follow-up: 3 months for 20/30 Adverse events monitored: no	1. Physician assessed effective rate for the 20 participants followed up for 3 months was 90% (18/20) 2. Marked effectiveness 80% (16/20), effectiveness 10% (2/20) and failure in 10% (2/20) 3. Significant increase in white cells/leukocytes and blood platelets following acupuncture 4. Significant increase in hematocrit and ESR decreased in the subgroup of 8/30 with anovulatory FUB. Do findings support conclusions: no – results presented as evidence of effectiveness for FUB and associated anaemia. Risk of bias and imprecision inherent in study design, from lack of validated outcomes, random error and observer bias not noted	a. *Change of blood picture: erythrocyte, leukocyte, haemoglobin, platelet indices* *b. Haematological indices: hematocrit, plasma viscosity, whole blood viscosity, fibrinogen, erythrocyte electrophoresis, erythrocyte sedimentation (ESR), K value of ESR equation* 1. Positive addition of "haemological" indices in an attempt to validate practitioner assessment

Study	Design and Methods	Participants	Interventions	Outcomes	Findings	Notes
Wang 1987 Primary dysmenorrhoea #16	Case-series ***Study QA Score:29% (7/24)*** Recruitment procedures: not stated Recruitment time period: 10 years Number of participants: 100 Homogeneity assessed: no Assessor blinding: not stated Lost to follow-up: not stated Follow-up long enough: yes Appropriate statistical analyses: NA Clarity of findings: partial Other evidence of quality control: no	Location: China Source population: not stated Inclusion criteria: implicit Exclusion criteria: not stated Characteristics of participants at intervention commencement: *Age*: Under 15 to over 26 years: 8/100 <15 years, 64/100 aged 16-20 , 22/100 aged 21-25, 6/100 aged >26 *Ethnicity*: not stated *Socio-econ. group*: not stated *Contraception*: not stated *Parity*: not stated *Obesity*: not stated *Severity*: not stated *Chronicity*: symptom duration from 1 to over 10 years: 1 to 5 years 72/100 6 to 10 years 24/100 Over 10 years 4/100 *Timing of dysmenorrhoea*: 44/100 pain prior to bleed 56/100 pain during bleed *Location of dysmenorrhoea*: 56/100 mid/low abdomen 44/100 bilateral low abd. *Menstrual flow*: 61/100 purplish dark 26/100 dark red 13/100 light red 31/100 profuse 69/100 scanty. *Accompanying symptoms*: nausea, vomiting, dizziness, and feeling of suffocation *Marital status*: 36/100 married, 64/100 single *Gynae. exams. of married women* : 7/36 anteversion of the uterus, 9/36 retroversion, 3/36 underdevelopment, 11/36 thickened annexa, 6/36 no abnormality	Electro-acupuncture, auricular-acupuncture and moxabustion according to TCM diagnosis ***Interventions QA Score: 37% (12.5/34)*** Setting: The People's Liberation Army 98 Hospital Practitioner details: not stated Style of acupuncture: implicit Rationale for treatments: syndrome patterns Points used: 2 formula per syndrome pattern Stagnation of Qi and Blood: (1) REN4, 17th V, LIV3 (2) REN6, LIV8, SP6 Retention of Cold: (1) DU4, SP9 (2) REN2, SP10, LI4 Deficiency of Qi and Blood: (1) REN6, BL18, ST36 (2) SP12, BL20, SP6 Needling details: needle type and depth of insertion not stated Needle stimulation: Stagnation of Qi and Blood: electro-acupuncture, frequency 150 per min. Retention of Cold and Deficiency of Qi and Blood not stated. Needle response: not stated Number of treatments: not stated Needle retention time: 30 minutes Frequency of treatments: 12 daily treatments per course beginning 8 days prior to onset of menstruation Co-interventions: (1) Auricular acupuncture with needles embedded: Stagnation of Qi and Blood: Liver and Endocrine. Retention of Cold: Kidney and Uterus. Deficiency of Qi and Blood: Spleen and Liver. Needle retention time not stated. (2) Moxabustion: Retention of Cold: Moxa on thin cakes of Radix Aconiti Praeparata at REN3, REN6, DU4. Deficiency of Qi and Blood: moxa sticks at ST36, SP6, REN6. Warming time and sensation not stated.	Physician assessed effective rate as sum of items A to C: *A. Cure*: complete relief of dysmenorrhoea and accompanying signs and symptoms for 6 months following treatment. *B. Marked effect*: marked relief. *C. Improvement*: some relief. *D. Failure:* no relief. Measures taken at baseline: not stated Measures taken at follow-up: not stated Evidence outcome measures are accurate: no Timing of follow-up: 6 months Adverse events monitored: no	1. Physician assessed total effective rate 94% (94/100) 2. Cure rate 54% 3. Therapeutic effect not related to severity of symptoms and signs Do findings support conclusions: no – results presented as evidence of effectiveness. Risk of bias and imprecision inherent in study design, from lack of validated outcomes, and observer bias not noted	

Study	Design and Methods	Participants	Interventions	Outcomes	Findings	Notes
Yuqin 1984 Dysmenorrhoea #17	Case-series ***Study QA Score:17% (4/24)*** Recruitment procedures: not stated Recruitment time period: not stated Number of participants: 49 Homogeneity addressed: partial Assessor blinding: not stated Lost to follow-up: not stated Follow-up long enough: partial Appropriate statistical analyses: NA Clarity of findings: Partial Other evidence of quality control: no	Location: China Source population: not stated Inclusion criteria: not stated Exclusion criteria: not stated Characteristics of participants at intervention commencement: *Age*: 14 -30 years 39/49 aged 14-20, 8/49 aged 21-25, 2/49 aged 26-30 *Ethnicity*: not stated *Socio-economic group*: not stated *Contraception*: not stated *Parity*: not stated *Obesity*: not stated *Severity*: not stated *Chronicity*: symptom duration 1-4 years: 19/49 duration 1-2 years, 7/49 > 2-4 years, 3/49 >4 years *Gynaecological exam*: 38/49 normal uterus, 6/49 mal-development, 4/49 slightly underdeveloped, 1/49 uterine fibroids *Type of dysmenorrhoea*: 48 primary and 1 secondary	Acupuncture, moxabustion and lifestyle advice according to TCM diagnosis ***Interventions QA Score: 36% (13/36)*** Setting: implicit – People's Hospital of Xuzhou District, Jiangsu Province Practitioner details: not stated Style of acupuncture: not stated Rationale for treatments: syndrome patterns Points used: Cold-Evil and Blood Stasis: REN6, ST29, SP10, BL32, SP6, BL23 Depression of Liver Energy and Stagnation of Vital Energy: REN4, BL32, ST25, P6, SP8, LIV3 Deficiency Cold of Vital Energy and Blood: REN6, ST29, BL23, SP20, SP10, ST36, SP6 Needling details: needle type and depth of insertion not stated Needle stimulation: Cold-Evil and Blood Stasis: mild reinforcing method, manipulated every 5mins. Depression of Liver Energy and Stagnation of Vital Energy: reinforcing method with manipulation. Deficiency Cold of Vital Energy and Blood: manipulation every 5mins. Needle response: not stated Number of treatments: 8 to 13, mean 10.5 Needle retention time: 20 to 30 minutes Frequency of treatments: every other day beginning 7-10 days prior to menstruation Co-interventions: (1) Moxabustion: Cold-Evil and Blood Stasis: at all points until skin reddened (2) Lifestyle advice: observe general relevant hygienic measures, avoid emotional disturbance, catching cold, raw food and cold food.	Physician assessed effective rate as sum of items A & B: *A. Complete cure*: 3 pain free cycles following treatment. *B. Marked improvement*: slight and bearable pain. *C. Failure*: no change. Measures taken at baseline: not stated Measures taken at follow-up: not stated Evidence outcome measures are accurate: no Timing of follow-up: 3 months Adverse events monitored: no	1. Physician assessed total effective rate 98% (48/49) 2. Complete cure 86% (42/49) 3. Effective rate for Cold-Evil and Blood Stasis 100% (31/31), complete cure 97% (30/31). 4. Depression of Liver Energy and Stagnation of Vital Energy 100% (12/12) (100%), complete cure 75% (9/12) 5. Deficiency Cold of Vital Energy and Blood 83% (5/6), complete cure 50% (3/6). 6. Fibroids were the probable cause of treatment failure 1/100. Do findings support conclusions: partial – no causal hypotheses made, but risk of bias and imprecision inherent in study design, from lack of validated outcomes, and observer bias not noted	

Study	Design and Methods	Participants	Interventions	Outcomes	Findings	Notes
Steinberger 1981 Primary dysmenorrhoea #18	Case-series ***Study QA Score:35% (8.5/24)*** Recruitment procedures: not stated Recruitment time period: 3 years Number of participants: 48 Homogeneity accounted for: no Assessor blinding: not stated Lost to follow-up: 4/48 – attrition not accounted for Follow-up long enough: yes Appropriate statistical analyses: NA Clarity of findings: partial Other evidence of quality control: no	Location: China – author is a Yugoslavian MD Source population: not stated Inclusion criteria: not stated Exclusion criteria: not stated Characteristics of participants at intervention commencement: *Age*: 19-42 years, mean 28 *Ethnicity*: not stated *Socio-economic group*: not stated *Contraception*: not stated *Parity*: not stated *Obesity*: not stated *Severity*: not stated *Chronicity*: symptom duration of 2-20 years *Main symptoms*: dysmenorrhoea with spasmodic pain in low abdomen, painful distension radiating to flanks, severe or persistent low back pain *Associated symptoms*: premenstrual tension, nausea, vomiting, headache, emotional disturbance.	Acupuncture formula and moxabustion where indicated ***Interventions QA Score: 66% (22.5/34)*** Setting: not stated Practitioner details: not stated Style of acupuncture: TCM Rationale for treatments: appropriate and easily located points that might be replicated by doctors inexperienced in acupuncture Points used: LI4, ST 36, SP 6, REN 4 Needling details: needle length 1.5 in, gauge 28 or 30. Depth of insertion not stated Needle stimulation and response: manipulation to obtain deqi Number of treatments: 5 to 10 Needle retention time: 20 to 30 minutes Frequency of treatments: daily for 5 days prior to menstruation, or occasionally during acute attack. Regimen repeated 2 months later if no effect experienced. Co-interventions: Moxabustion at ST 36 for nausea and vomiting, and at REN 4 for severe pain. Warming time and response not stated.	Physician assessed effective rate as sum of items A & B: *Good*: no pain or discomfort for at least 6 months post treatment, *Satisfactory*: less pain and discomfort. *Poor*: no relief after second course of treatment. Measures taken at baseline: not stated Measures taken at follow-up: not stated Evidence outcome measures are accurate: no Timing of follow-up: 6 months Adverse events monitored: no	1. Physician assessed total effective rate 85% (40/48) 2. Good response 58% (28/48). Do findings support conclusions: partial – no causal hypotheses made, but risk of bias and imprecision inherent in study design, from lack of validated outcomes, and observer bias not noted	

ANNEX B: Acupuncture for Menorrhagia Review: Table of excluded studies

Study	Reason for exclusion
Ji 1999	This cohort study was carried out in China and compared electro-acupuncture formulae prescribed according to syndrome patterns to modern medicine (western drug therapy) in the treatment of 98 women with infertility due to irregular menstruation (anovulation). The outcomes of interest were ovulation and pregnancy, not the regulation of menstruation as evidenced by normalisation of menstrual flow and amelioration of pain and discomfort, so the trial was excluded. Design & methods: cohort study, source population and recruitment procedures not stated, recruitment period 6 years and 10 months, baseline comparability not stated, assessor blinding not stated, loss to follow-up not stated, intention to treat not stated. Appropriate statistical analyses with t-test and chi squared, but lack of simple outcome data denies alternative analyses. Participants: 98, 68 acupuncture group, 30 modern medicine group. Inclusion criteria: infertility >1year with amenorrhoea or irregular menstruation, mono-phase basal body temperature, abnormal serum follicle stimulating hormone (FSH), luteinizing hormone (LH), prolactin (PRL), estradiol (E2) and (P) progesterone values, endometrial glandular secretion not good. Aged 23-40. Infertility for 1-5 years. Previous modern medicine ineffective for 36/68 in acupuncture group. Salpingitis 21/68 in acupuncture group, 1/30 in modern medicine group. Interventions: *Acupuncture rationale: TCM syndrome patterns*: Main points: DU20, REN4, ST29, Zigong, SP6. Adjuvant points: Kidney Deficiency BL23, DU4. Reinforcing method, deqi obtained. Liver Stagnation BL18, BL19, LIV3. Reducing method, deqi obtained. Blood Stasis REN2, BL17. Reducing method, deqi obtained. Phlegm-Dampness SP9, ST40. Reducing method, deqi obtained. Electro-stimulation with bearable intensity (no further information). Duration/dosage: 30 min treatment, 10 daily/course (number of courses not stated). *Modern medicine*: clomiphene, chorionic gonadotropin, stilbestrol and progesterone. Duration/dosage: for 5-6 menstrual cycles clomiphene 50mg daily for 5 days from 5th day of menstruation with chorionic gonadotropin 5000i.u. injected intramuscularly for 3-4 days. From 10th day of cycle stilbestrol 0.5mg daily for 10 days. If no ovulation, progesterone 20mg intramuscularly injected from day 20 daily for 5 days to cause withdrawal bleed. Outcomes: Ovulation using basal body temperature readings and endometrial biopsy of corpus luteum late stage, and changes of serum FSH, LH, E2 and PRL tested and analysed using a radio-immunoassay box, pregnancy rate, natural abortion and ovary excitation. Evidence outcome measures are accurate is partial due to limited information. Baseline measures: BBT, biopsy or corpus luteum and FSH, LH, E2 and PRL values. Follow-up at 6 months: FSH, LH, E2 and PRL values, pregnancy rate, natural aborption and ovary excitation. Monitoring of adverse events not stated. Findings: Total ovulation rate 65/68 (79.4%) acupuncture group, 28/30 (83.3%) modern medicine group. Pregnancy rate 47/68 (69.1%) acupuncture group, 12/30 (40.1%) modern medicine group. Ovary excitation 0/68 acupuncture group, 8/30 modern medicine group. Natural abortion 0/68 in acupuncture group, 2/30 in modern medicine group. Hormone values FSH, LH and E2 increased and PRL decreased in both groups (normal values not given). Risk of bias due to confounding not noted by author.
Sha 2000	This RCT was carried out in China and compared acupuncture, moxabustion and cupping according to syndrome patterns to drug therapy in the treatment of 151 patients with premature ovarian failure. The outcomes of interest were ovulation and pregnancy, not the regulation of menstruation as evidenced by normalisation of menstrual flow and amelioration of pain and discomfort, so the trial was excluded. Design & methods: RCT, source population and recruitment procedures not stated, recruitment time-period not stated, baseline comparability not stated, assessor blinding not stated, loss to follow-up not stated, intention to treat analysis not stated. Statistical analyses do not give actual probability values and provide few absolute numbers. Participants: 76 acupuncture group (AG), 75 drug therapy (DTG). Inclusion criteria: premature ovarian failure (PRL normal, FSH>40*mg*/L. Clinical symptoms present were disorder of menstruation followed by amenorrhoea, regular followed by irregular menstruation, sudden onset, amenorrhoea, hectic fever and sexual dysfunction or atrophy of edeatrophia. Patient history might include abnormal chromosome, viral or bacterial infection, immunic disease or hereditary family history. Age range 26-40 years, mean 35. Duration of symptoms 1-9 years, mean 5-6 years. TCM diagnosis: Liver and Kidney Yin Deficiency 103 (AG 55/76, DTG 48/75), Spleen and Kidney Yang Deficiency 48 (AG 21/76, DTG 27/75). Interventions: setting – Shangdong Provincial Hospital. (1) *Acupuncture rationale TCM syndrome patterns.* Main points: REN4, REN3, KID12, Zigong, BL23, Jiaji, thoracic 5th, lumbar 4th Plus relevant adjuvant points:*Liver and Kidney Yin Deficiency*: SP6, SP9, BL18, HT6, KID7, *Spleen and Kidney Yang Deficiency*: DU4, SP8. Reinforcing method, deqi obtained. Up to 6 courses of 20 treatments. Needle retention time 20 minutes. Frequency of treatments: 5 to 7 days between courses. Co-interventions: Moxibustion then cupping for 5 to 10 minutes at BL20, BL23 and BL32 for cases of *Spleen and Kidney Yang Deficiency. (2) Drug therapy:* Clomiphene and stilbestrol combined Treatment regimen: Clomiphene 50mg/day for 5 days, stilbestrol 0.5-1.0mg for 20 days. Up to 6 courses with 5-7 days between courses. Rationale/sources that justify control: not stated. Outcomes: Physician assessed effective rate using clinical records and radioimmunological study of FSH, LH and E2 hormone values. The effective rate is sum of A and B. A= *Markedly effective*: amenorrhoea, irregular menstruation, hectic fever and sweating, dizziness, restlessness, insomnia, sexual dysfunction, dryness of vagina and atrophy of genitals disappeared or markedly relieved. FSH ≤40mg/L, menstruation and ovulatory functions normal or pregnancy. B = *Improved*: improvement in symptoms, normal menstruation and hormone values. C. *Failed*: no dramatic change. Measures taken at baseline and follow-up: PRL, FSH and LH values. Evidence outcome measures are accurate: no. Timing of follow-up: post-treatment, 30 and 90 days. Adverse events monitored: no Findings: 1. *Effective rate*: AG 89.5% (68/76), DTG 64% (48/75). 2. *Markedly effective rate*: AG 68.4% (52/76), CG 28% (21/75). 3. *Hormone Values:* FSH, LH and E2 changed in both groups, with changes maintained post treatment in the acupuncture group (change in relation to normal values not given). Risk of bias and imprecision from inadequate randomisation, prognostic variables, lack of validated outcome measures and observer bias not noted by author
Griffiths 2000	This study was excluded because it was a case report. Written by a nurse-acupuncturist, it described the positive response of an 18 year-old patient with primary dysmenorrhoea treated by traditional acupuncture and Chinese herbs at a clinic in Australia over a period of 6 months.
Hu 1999	This study was excluded because it was a case report. Written by a Chinese acupuncturist, it described the positive response of a 21 year-old with primary dysmenorrhoea treated by traditional acupuncture over six cycles.

Chapter Three

The York and Selby ACUMEN Exploratory Trial:

An exploratory trial to assess the feasibility of a full-scale pragmatic randomised trial to evaluate the clinical and economic benefits of offering acupuncture to patients with menorrhagia in primary care

1. Acronym

ACUMEN (ACUpuncture for MENorrhagia exploratory trial)

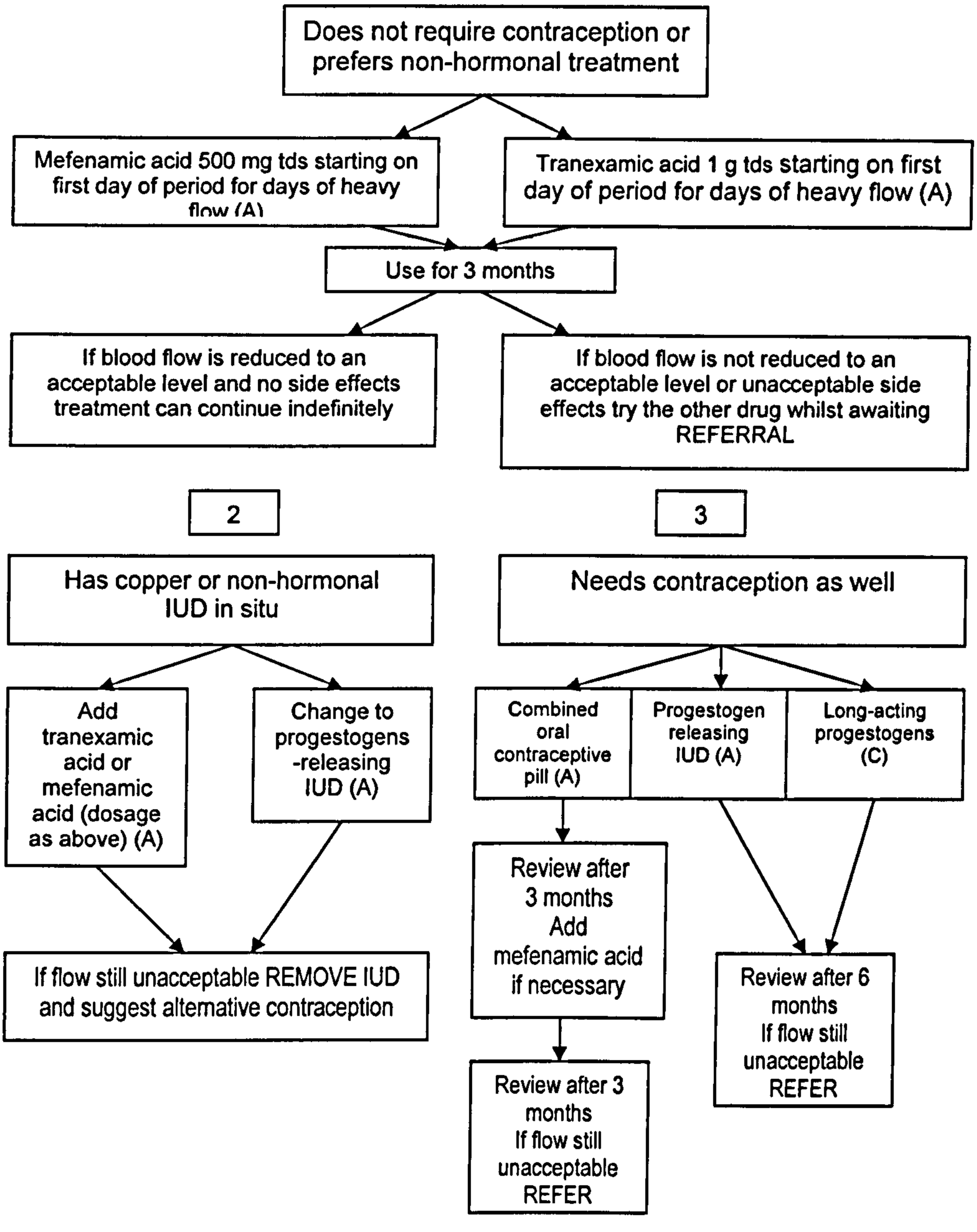
2. Introduction and rationale

2.1 A potential role for acupuncture

Chapter One, Section 1.2 of the thesis describes the drug therapies currently available to women with menorrhagia in primary care *[Figure 4]*and the limitations of these treatments, where some women find them ineffective, unacceptable because of side-effects, or inappropriate due to concurrent medical conditions (Irvine et al 1999; Bonar & Sheppard 1997; Coulter et al 1995; Lumsden & Smith 1992; Nilsson & Rybo 1967); with other women having a preference for surgery over long term drug therapies (Marshall 1998; Coulter et al 1995), or preferring therapies that do not involve drugs or surgery (Longridge nee Gamon 2000; Scambler & Scambler 1985). This chapter also outlined the therapeutic approach of traditional acupuncture *[Chapter Zero, 1.4]*, a natural, low-tech therapy that is available to women with menorrhagia in the private healthcare sector from members of the British Acupuncture Council (Wadlow & Peringer 1998), and considered acupuncture might usefully add to the therapeutic options currently available to women in the NHS, so enabling more women to effectively manage their symptoms within primary care and avoid, rather than merely delay, referral for what has been termed a *"surgical solution"* (Lilford 1997).

The systematic review of acupuncture for menorrhagia presented in Chapter Two of the thesis found just two studies that had sought to evaluate acupuncture in the treatment of menorrhagia or excessive menstrual loss. One involved the treatment of 50 (Zhang 1994) and the other 30 women (Wangcheng 1988). The positive findings from these studies are purely suggestive as both were case series studies, and in addition to the bias and imprecision

Figure 4: **Medical management of the complaint of menorrhagia** from *"Guideline: The Initial Management of Menorrhagia"* published by the Royal College of Obstetrics and Gynaecology, UK www.rcog.org.uk/guidelines/menorrhagia.html 01/02/00



Whilst oral luteal phase progestogens are ineffective in reducing menstrual blood loss (A), intrauterine progesterones are effective (A)

A = based on randomised trials
B = based on other robust experimental or observational studies
C = based on more limited evidence but the advice relies on expert opinion and has the endorsement of respected authorities

Note: This guideline was in date at the time of the ACUMEN Study

inherent in the study design, both suffered from a lack of methodological rigour whereby recruitment procedures and interventions were poorly defined and validated or credible outcomes were lacking. And the four studies that concerned the treatment of amenorrhoea or irregular menstruation suffered from similar problems (Ma 1999; Tureanu 1994; Mo 1993; Liu 1992), in spite of one being a controlled trial (Ma 1999). The majority of the evidence available concerned the treatment of primary dysmenorrhoea (12 of 18 studies), and most were case series studies (10/12); but two pilot trials indicate a possible beneficial effect from acupuncture (Helms 1987; Thomas 1985), although a lack of methodological rigour, such as inadequate randomisation procedures and inappropriate statistical tests, small sample sizes and evidence of confounding from carry-over effects diminish the potential usefulness of results. And yet, the interest in the potential for acupuncture in the treatment of gynaecological conditions is growing. At the time of ACUMEN, in Sweden Stener-Victorin (2000) published evidence of the effectiveness of electro-acupuncture for the treatment of anovulation from polycystic ovary syndrome, and a study of the effectiveness of acupuncture for primary dysmenorrhoea was being conducted (Smith 2000). Arguably the rationale for this research resides in the safety of acupuncture *[Chapter One, 1.5.1]* (MacPherson et al 2001; Ernst et al 2001; MacPherson et al. 2004) alongside the continued need for more supportive and less invasive treatment options.

Moreover, past studies have found high levels of satisfaction with complementary and alternative medicine (CAM), both amongst private and NHS users (Luff & Thomas 2000; Gould & MacPherson 2001; Paterson & Britten 1999; Cassidy 1998a; Cassidy 1998b; Kelner & Wellman 1997; Vincent & Furnham 1996). Overall, patients have been found to have a strong, pragmatic, disease-management motivation for CAM use, to develop an increasing commitment to CAM based on positive treatment experiences, and to experience beneficial patient-practitioner relationships. A qualitative study by Luff & Thomas (2000) involving 49 NHS patients, 21 of whom had received acupuncture, also revealed how a commitment to treatment seems to develop as a direct result of user satisfaction with treatment, rather than because of prior belief about complementary therapies, or any commitment arising from financial investments. Moreover, a survey by Gould & MacPherson (2001) involving 72 patients receiving treatment privately from members of the British Acupuncture Council, suggests that patients can experience a wide-range of benefits in response to acupuncture, that they like the holistic focus of treatment, and are willing to comply with lifestyle advice to aid a

return to, or maintain, good health. And in the US survey of patients consulting acupuncturists, the majority reported high levels of satisfaction with their care, of whom 17.4% (100/575) had received acupuncture for gynaecological conditions (Cassidy 1998a & 1998b). Indeed, current estimates show that one in five people in the UK use some form of complementary and alternative medicine (CAM) (Ernst & White 2000). Acupuncture is one of the most popular CAM modalities, and even almost 10 years ago, at least 6% of those patients privately consulting traditional acupuncture practitioners in the UK received treatment for gynaecological and obstetric conditions (Wadlow & Peringer 1998).

2.2 The need for an exploratory trial

Research is needed to develop the evidence base for traditional acupuncture in the treatment of menorrhagia. The gold standard for evaluating medical treatments is the randomised controlled trial (RCT). A pragmatic RCT or other rigorous study to answer questions about the relative clinical and cost effectiveness of different treatment options for menorrhagia, including acupuncture, in normal primary care would enhance the current evidence base. Such research would be entirely consistent with the House of Lords (2000) recommendations. In particular it would aid practical decisions about the use of acupuncture in the NHS; provide practitioners of traditional acupuncture with valuable information about its' potential role in the treatment of menorrhagia; and enable women to decide whether to invest in private acupuncture for symptoms of excessive menstrual bleeding. A necessary precursor to a full-scale clinical trial is a pilot or exploratory trial to test and develop a protocol for the rigorous evaluation of the clinical and economic benefits of offering acupuncture to patients with menorrhagia assessed as suitable for management in primary care.

3. The aims and objectives of the ACUMEN exploratory trial

Aim:

- To study the justification for, and feasibility of, a full-scale pragmatic randomised trial to evaluate the clinical and economic benefits of offering acupuncture to patients with menorrhagia in primary care.

Objectives:

- To test GP recruitment procedures,
- To test patient recruitment procedures,
- To estimate the patient recruitment rate,
- To test the referral procedures to acupuncture treatment,
- To assess the acupuncture treatment protocol,
- To assess the comparator treatment protocol,
- To investigate compliance with acupuncture treatment,
- To choose outcome measures for menorrhagia,
- To assess the appropriateness of these outcome measures for a full trial,
- To estimate loss to follow-up, and
- To estimate the likely effect size from acupuncture, and thus the sample size required for a full trial.

4. Methods - protocol

4.1 Explanatory and pragmatic randomised trials

Randomised controlled trials are considered the most rigorous way of determining whether a cause-effect relation exists between treatment and outcome and for assessing the cost-effectiveness of a treatment (Sibbald & Roland 1998). They are defined as clinical experiments in which patients receive treatments which are allocated to them according to some method of randomisation (Pocock 1983). They are conducted when there is genuine doubt about whether one treatment approach is better than another. The individuals who might benefit are randomly allocated to receive the experimental intervention or not; the latter form the control group and receive a placebo, no treatment or usual care. The type of control used is indicative of the type or randomised trial conducted and the nature of the question addressed *[Table 19]*.

Fastidious or explanatory randomised trials involve a placebo to isolate any effects from the participant or caregiver knowing the type of treatment received – the psychosomatic effects – in order to equalise these non-specific effects between groups and study the 'true', 'specific effects' of the treatment (Schwartz & Lellouch 1967). A pragmatic trial uses no treatment, "usual care" or "best practice" as the control and instead includes the non-specific effects

within the specific effects of the treatment. The more complex the delivery of the intervention, the more difficult it is to isolate the specific and non-specific effects of treatment. Thus, whilst drug therapies are typically tested using the explanatory model, it is impossible to compare two types of psychotherapy without including any placebo effects within those of the treatments.

Table 19: Differences between fastidious and pragmatic trials

	Fastidious trials	Pragmatic trials
Objective	To draw *conclusions* about therapies by testing defined scientific hypotheses	To make *decisions* between therapies in clinical practice
Experimental conditions	Laboratory conditions	Normal clinical practice
Definition of therapies	1 Rigid 2 Equalised -- therapies are defined to achieve the *same* placebo effect	1 Flexible 2 Optimal -- each therapy is defined to make the *best* of any placebo effect
Definition of patients	1 Patients eligible for the trial (ie for all therapies) are strictly defined a priori and may be redefined a posteriori 2 Patients who withdraw from therapy are *withdrawn* from the trial	1 Patients eligible for the trial (ie for all therapies) are flexibly but irrevocably defined a priori 2 Patients who withdraw from therapy *remain* in the trial for analysis
Number of criteria	Single or multiple	Single criterion, if necessary by combining multiple criteria
Method of analysis	Traditional significance test for each hypothesis, but no formal relationship between significance tests	Select therapy that gives best weighted criterion (no formal significance test)
Number of patients	Traditional calculation based on *type 1 and type 2 errors* for each hypothesis; the total number of patients should be the maximum of these calculations	Calculation based on the weighted criterion and on *type 3 error*, ie concluding that therapy A is superior to therapy B when the opposite is true

Sources:

Schwartz D & Lellouch J.1967 Explanatory and pragmatic attitudes in clinical trials. Journal of Chronic Diseases, 20, 637-644.

Russell IT (1983). The evaluation of computerised tomography: a review of research methods. In Economic and medical evaluation of health care technologies, eds Culyer AJ, Horsberger B. Berlin: Springer-Verlag.

Similarly, double-blinding is impossible in acupuncture trials as the caregiver must consciously perform either 'real' or 'sham' acupuncture. And because the treatment also needs to be fastidiously applied, for traditional acupuncture, this has usually meant using the same acupoints for all patients rather than individualised prescriptions according to theory, as well as omitting other elements of usual practice, such as moxibustion or dietary therapy, so as not to contaminate the results. This rigid control is then extended to the study patients who are required to be as homogenous as possible to further minimise error through variance and better enable conclusions to be drawn. Thus the explanatory model is best able to ask very precise theoretical questions about the effect of a specific element of a complex intervention, such as acupuncture, when conducted under laboratory like conditions. However, the testing of clearly defined scientific hypotheses rarely informs practice, because whilst internal validity is often high, external validity and the generalisability of results is severely compromised.

The pragmatic method, by contrast, includes those patients that would *normally* be eligible for treatment, and always provides an answer to the practical question of which treatment is better when administered under normal health care conditions, with each therapy making the best of any placebo effect. The principal method for ensuring unbiased outcome assessment is proper randomisation (Pocock 1983; Campbell & Machin 1999).

The main purpose of randomisation is to avoid bias by distributing the characteristics of patients that may influence outcome randomly between treatment groups so that any difference in outcome can be explained only by treatment (Roberts & Torgerson 1998). These characteristics might be demographic ones such as age or prognostic factors such as clinical history or disease severity. Simple, elegant randomisation where unpredictable, reliable allocation is achieved involves either a table of random numbers or a computer random number generator. To ensure the desired treatment group sizes are achieved, especially where sample sizes are less than 200, a system of random blocks is usually used. For example, if blocks of four are used there are six sequences by which treatments A and B can be equally allocated: AABB; ABAB; ABBA; BAAB; BABA and BBAA. However, randomly varying the block sizes and using larger block sizes, particularly in unblinded trials, better ensures the generation of an unpredictable randomised allocation sequence and concealment of that sequence from those responsible for recruiting participants until assignment occurs (Pocock 1983).

However, imbalances in prognostic characteristics can still arise by chance and bias the analysis of outcome. Stratified blocks or minimisation can achieve balance (Roberts & Torgerson 1998). Stratification according to one or two factors is a precise but complex method that is particularly useful in small trials where it can avert severe imbalances on prognostic factors. In larger trials, where participants per group exceed 50, the control exerted to achieve stratification compromises unpredictability and is often unnecessary where randomisation naturally achieves balance. Minimisation seeks to achieve balance on a set of prognostic factors, although not for each combination (Roberts & Torgerson 1998; Altman 1991). Its supporters describe it as a more easily applicable system that can generate similar groups even in small trials (Treasure & MacRae 1998), whilst its detractors are most aware of the increased risk of trial participants being able to guess upcoming assignments (Pocock 1983).

Whatever the method of allocation, allocation concealment is essential if selection bias is to be eliminated (who gets into the trial and the treatment they are assigned) and unbiased comparison groups created (Chalmers 2001). With all approaches, the people who generated the allocations scheme should not be involved in ascertaining eligibility, administering treatment, or assessing outcome (Russell 1999). An increasingly common procedure is to use a central telephone randomisation service, such as the one run by the Department of Health Sciences, University of York. Here patient details are supplied, eligibility confirmed and the patient entered into the trial before treatment allocation is divulged (and it may still be blinded). If external help is not available the only other system that provides a plausible defence against allocation bias is to enclose assignments in serially numbered opaque, sealed envelopes. This method is not immune to corruption however.

Once proper randomisation is assured, and assuming the validity of the interventions, unbiased assessment of outcome rests upon the appropriateness and reliability of the measures of outcome used, whether or not participants are excluded from the analysis and the statistical tests used. In an explanatory trial, a single validated physiological or biological outcome is often chosen (Roland & Torgerson 1998). In pragmatic trials, a single outcome measure is usually inadequate for clinicians and other healthcare decision makers to weigh up the risks, costs, and benefits of a given intervention. A common approach is to use both a general-health and disease specific measure within one trial.

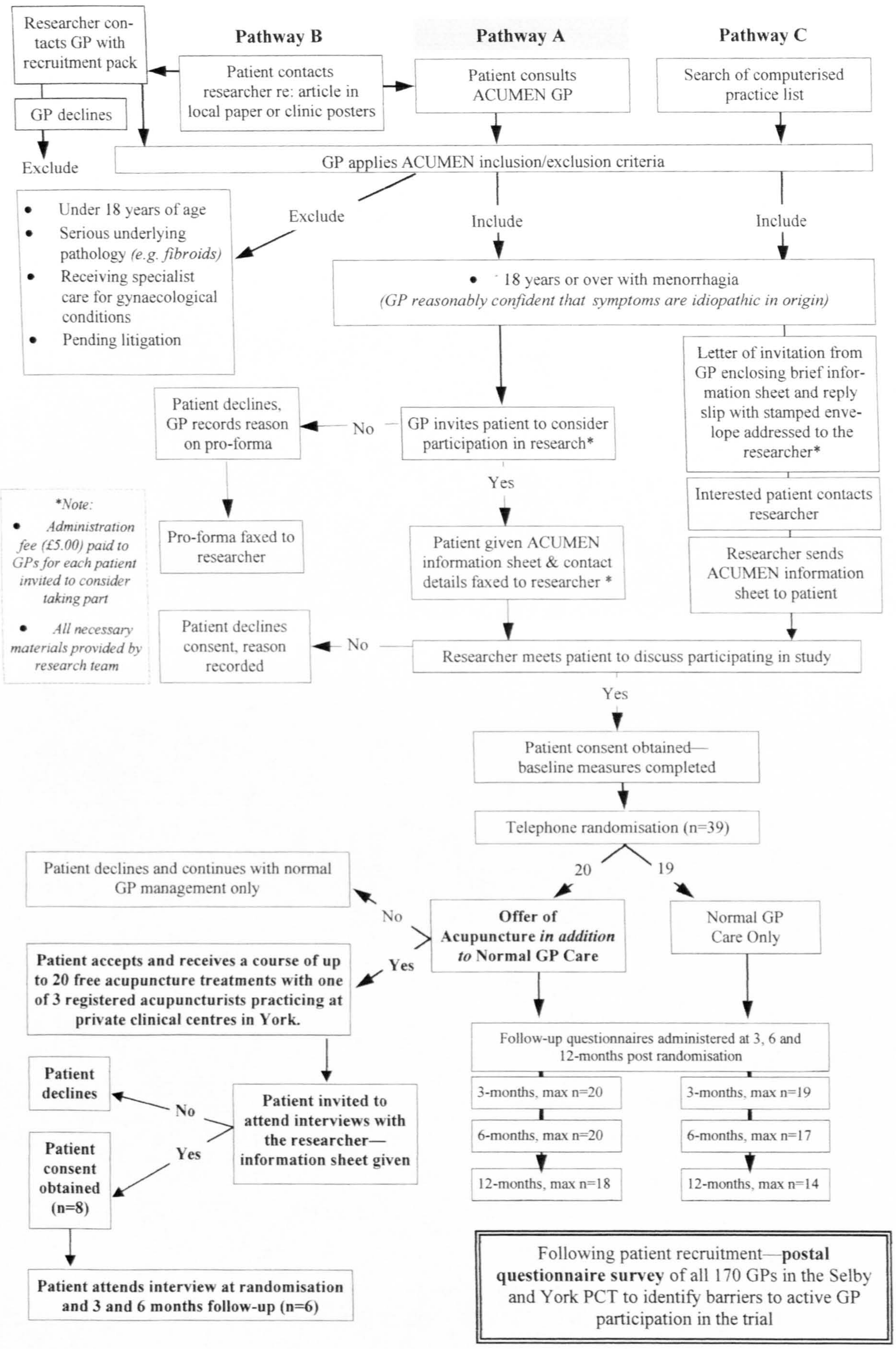
With regards the analysis, in explanatory trials, where drawing a conclusion about efficacy is the aim, patients not receiving the treatment as designed are excluded to avoid introducing bias and to minimise Type2 Error (wrongly conclude treatments A & B are equivalent) (Campbell & Machin 1999; Schwartz & Lellouch 1967). Conversely, in pragmatic trials, to ensure generalisability and minimise Type1 Error (wrongly conclude treatments A & B differ), 'intention-to-treat analysis' is essential. Here all participants are followed up irrespective of whether they actually receive the intervention offered. Thus 'intention-to-treat analysis' works to maintain the advantages of random allocation, which may be lost if patients are excluded from the analysis, if they withdraw from the trial, or do not comply with treatment.

Traditionally, statistical comparisons in randomised trials tend to involve either a comparison of follow-up scores or a comparison of change scores, where a change score is calculated by subtracting the follow-up score from the baseline score. However, in 2001 Vickers and Altman showed these methods to introduce bias and determined that a better, more efficient approach was to use analysis of covariance (ANCOVA). This test adjusts each patient's follow-up score for his of her baseline score of the outcome measure in question, and has the advantage of being unaffected by baseline differences. This is important, as if by chance baseline scores are worse in the treatment group, the treatment effect will be underestimated by a follow-up score analysis and over estimated by looking at change scores, because of regression to the mean. By contrast ANCOVA gives the same answer whether or not there is baseline imbalance and has greater statistical power to detect a treatment effect than other methods.

4.2 The ACUMEN exploratory trial design

The practical aim of the ACUMEN trial dictated a pragmatic approach. The basic trial design was that of a pragmatic randomised controlled trial *[Table 19]* (Schwartz & Lellouch, 1967) involving 39 patients. The trial was set within the geographical boundaries of the Selby and

Fig 5: ACUMEN exploratory trial with nested qualitative study and survey - emergent trial design



York Primary Care Trust. Those 19 patients randomised to the control or comparator arm received usual care only from their GPs; but their use of private treatment outside the trial was monitored at follow-up. The 20 patients randomised to the experimental arm, *in addition to* usual care from their GP, were *offered* referral for traditional acupuncture with one of three professional acupuncture practitioners at one of three private clinical centres in York.

Thus the study was designed to allow a comparison of different treatment options, or packages of care – Care Package A being the therapies usually available to women with menorrhagia in primary care; Care Package B being the addition of acupuncture to the range of therapies currently available through their GP. In this way, the design took into consideration a number of key issues. First, by seeking to evaluate the effects of usual practice and so allowing the tailoring of the intervention to the individual and the development of a beneficial therapeutic relationship, it was congruent with the principles of traditional acupuncture.

Second, by *offering* acupuncture within normal NHS management with uptake depending on patients, it recognised the issue of patient preference and maximised the placebo effect of patient preference – with patients able to choose from their allocated package of care the treatment most attractive to them (Fitter & Thomas 1997; Thomas et al. 1999). Treatment preferences were ascertained prior to total randomisation in order that patients' enthusiasm or disappointment about their treatment assignment could be taken into consideration in the analysis. This method was advised by Torgerson and colleagues (1996), who found there to be important differences between those with a preference and those indifferent to treatments in their trial of exercise for back pain compared to usual GP care. Whereby if patients had been either allocated to their treatment of choice or randomised when indifferent in accordance with the Brewin-Bradley design (Brewin & Bradley 1989), they would have been unable to determine whether greater benefits in the exercise group were due to enhanced motivation or the fact they had on average had back pain for a shorter length of time than those in the control group or perhaps both (Torgerson et al. 1996).

Third, by providing patients with acupuncture *in addition to* usual care, the design recognised that usual practice allows the provision of acupuncture alongside prescribed drug therapies, which may be reduced in line with benefits from acupuncture (Paterson & Britten 2000).

Fourth and finally, it sought to minimise patient attrition, whether or not acupuncture was chosen.

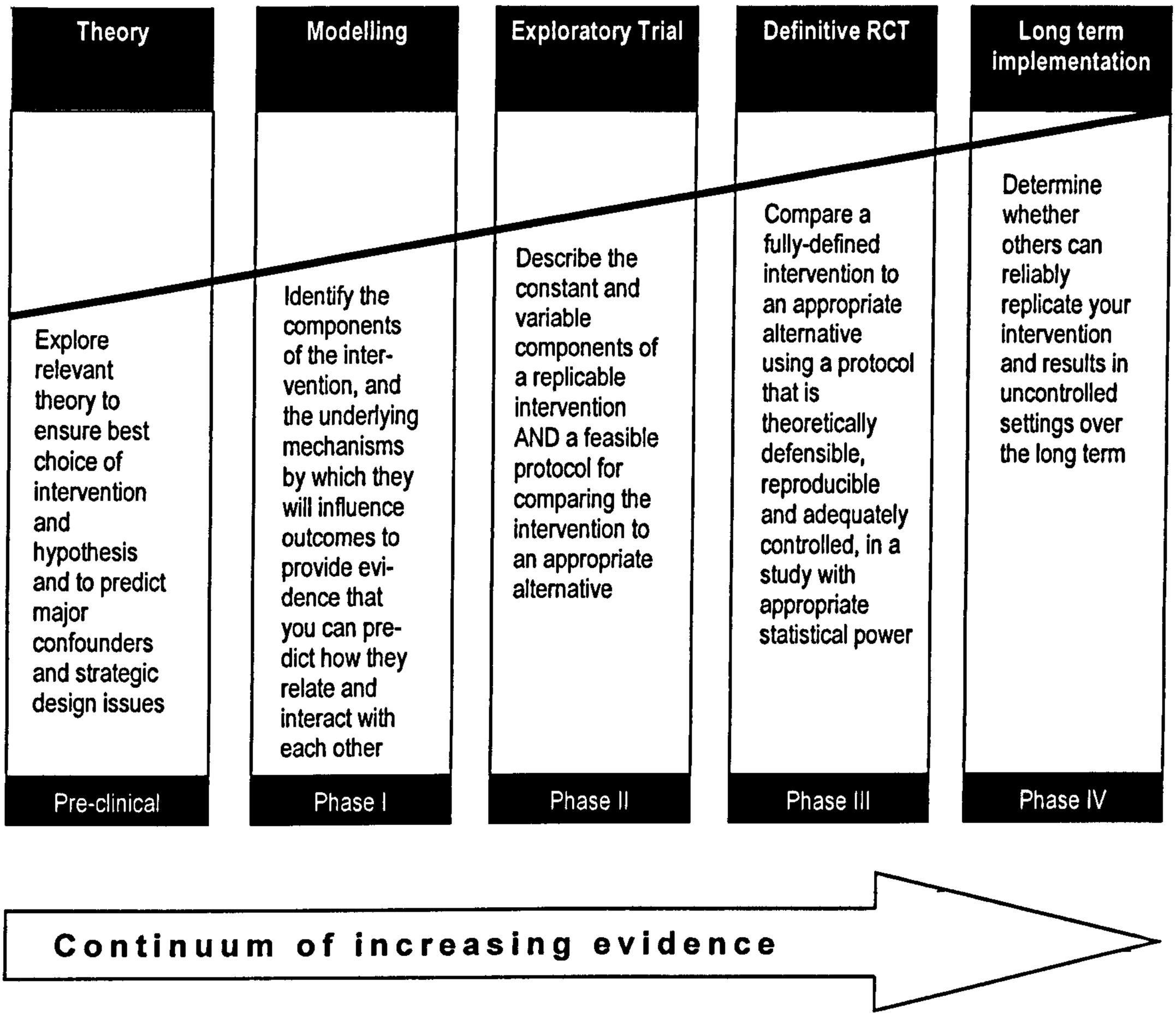
The ACUMEN exploratory trial therefore built on a design successfully implemented by the Medical Care Research Unit, University of Sheffield, and Foundation for Traditional Chinese Medicine (FTCM), York – two of the research centres involved in ACUMEN – for YACBAC, the trial of acupuncture for low back pain in York funded by the NHS Executive (Thomas et al. 1999). A requirement for trials funded by the FTCM is that they allow for a complex acupuncture intervention that closely mirrors the care usually provided by members of the British Acupuncture Council. The key differences between the YACBAC and ACUMEN designs were the additional recruitment pathways employed in ACUMEN, the nested qualitative investigation and the complementary postal questionnaire survey *[Figure 5]*.

4.2.1 The evaluation of complex interventions to improve health

The ACUMEN exploratory trial also used the framework for the development and evaluation of complex interventions to improve health proposed by the Medical Research Council (2000) in a discussion document published in April 2000; it combines elements of both phases 1 and 2 *[Figure 6]*. Traditional acupuncture meets the criteria for a "complex intervention": it responds to the unique way in which disease manifests itself in individuals; it tailors treatment (needles, moxibustion, cupping and massage) to the individual; it emphasises education and self-care according to aetiology, whereby advice on diet, rest and exercise is offered as appropriate; and it recognises the quality of the patient-practitioner relationship as an important factor in bringing about health and healing (MacPherson & Kaptchuk 1997) *[Chapter One, 1.4]*. Thus traditional acupuncture is *"built up from a number of components, which may act both independently and inter-dependently"* (MRC 2000).

The prospective longitudinal qualitative study interviewed six of the 20 women accepting the offer of acupuncture in addition to usual GP care within the trial. The issue of "treatment acceptability" was explored before, during and following treatment, and valuable information about the *"complex package of 'active ingredients'"* (MRC 2000) delivered within

Figure 6: **Framework for trials of complex interventions –** reproduced from *" A framework for development and evaluation of RCT's for complex interventions to improve health: a discussion document"* drafted by members of the MRC Health Services and Public Health Research Board, April 2000



ACUMEN elicited. The survey questionnaire, sent to all 170 GPs within the Selby and York PCT, aimed to identify the barriers to active GP participation in the trial. These studies, reported in Chapters Four and Five of the thesis, provide important additional information on the feasibility of, and justification for, a full-scale trial, particularly GP and patient recruitment strategies, the content of the patient information sheet, and delivery of the intervention. Thus the qualitative study and survey address issues in Phase 1 of the MRC framework, namely *"unravelling and distinguishing the key components in a complex intervention"* and *"identifying beliefs, behaviours or organisational factors that make an intervention more or less plausible"*. The trial itself addresses issues in Phase 2, namely testing the *"feasibility of key components of a larger trial, such as recruitment, randomisation and measurement of outcome"* and providing *"evidence of intervention effects for the purposes of calculating the power of a larger trial."*

4.3 Recruitment of participants

4.3.1 Recruitment of GPs

The ACUMEN research team had predicted that the recruitment target of 40 patients would be achieved over the course of five months with the support of 10 GPs. However, after three months just six patients had been recruited to the study compared with a target of 24. As a result the recruitment period was extended by six months and the researcher continued to ask for GP support *[Appendices 7 & 8]*. After five months, 32 GPs (from 15 practices) in the Selby and York Primary Care Trust had formally offered their support. And by the end of the recruitment period, another 11 GPs had facilitated recruitment for 16 patients who first contacted the researcher to express an interest in taking part. Thus a total of 43 GPs were involved in patient recruitment.

Newsletters were sent to all GPs in the trust via the internal mail system. These sought to raise the profile of the study, encourage more active participation, update them on the trial's progress, and address pertinent issues. A total of four newsletters were issued over the course of ACUMEN *[Appendices 9 to 12]*. These were dated September 2001, November 2001, March 2002 and March 2003. The first issue focused upon answering questions that were commonly raised during meetings with GPs, for example the rationale for conducting a pragmatic as opposed to explanatory RCT, and changes to normal practice imposed by the

trial. The second issue informed GPs of the success of patient-centred recruitment strategies, the third informed them about the forthcoming ACUMEN GP survey to understand referral patterns *[Chapter Four]*, and the fourth announced the winner of the GP Survey Prize Draw.

4.3.2 Recruitment of patients

Table 20: Patient inclusion and exclusion criteria

Inclusion criteria *Women were eligible if:* • They were aged 18 years and over, and • They had menorrhagia assessed as suitable for management in primary care
Exclusion criteria *Women were not eligible if:* • They had bleeding arising from the genital tract in the presence of suspected or serious underlying pathology such as carcinoma, endometriosis or fibroids, where referral to the specialist is the option of choice, • They were currently receiving specialist medical care for gynaecological conditions, • They had health conditions for which needling is contraindicated (e.g. haemophilia and susceptibility to lymphoedema), • They were pending litigation, • They could not read and write fluently in English.

This study involved women of 18 years and over with menorrhagia assessed as suitable for management in primary care by their GP *[Table 20]*. Thirty-nine patients out of a target of 40 were recruited to the trial via one planned and two emergent recruitment pathways and over an extended 11-month recruitment period *[Figure 5]*. Chapter Five of the thesis discusses recruitment in detail in relation to the GP survey that came about in order to understand referral patterns:

Patient recruitment pathway A

GPs were asked to identify suitable patients during a consultation and invite them to consider participation in a trial that may involve the offer of acupuncture, with the help of posters displayed in the surgery and waiting room *[Appendix 13]*. Thus this was similar to the recruitment strategy that had been successfully employed in the YACBAC trial, where 43 GPs

recruited 241 patients over a period of 18-months (Thorpe & Bell-Syer 2002). GPs provided patients with a copy of the information sheet *[Appendix 14]* and faxed a pro-forma *[Appendix 15]* to the candidate with the patient's age and duration of symptoms of menorrhagia. For patients interested in participating, GPs obtained their informed consent to transfer their contact details (name, address and telephone number) to the candidate. Where the patient declined to consider trial participation, GPs were asked to record their reason.

However, after three months of a planned five-month recruitment period, only six of a target of 24 patients had been recruited, and just one of these six referrals had been initiated by the GP within a consultation *[pathway A]*. The other five had instead been initiated by the patient, either in response to GP surgery posters (3/6), or on the recommendation of an ACUMEN patient or acupuncture practitioner (2/6), by first contacting the candidate. Thus, the trial team extended the recruitment period by six months and decided to pilot other patient-centred strategies:

Patient recruitment pathway B

First, a health care journalist from a local newspaper was invited to write an article about the study, and posters similar to those in GP surgeries were displayed in the three participating acupuncture clinics. This article involved interviews with the candidate, the three women that had been randomised to the offer of acupuncture, and one of the ACUMEN practitioners, and the ACUMEN practitioner and one of the women also consented to photographs. The article concluded by inviting women with heavy periods to contact the researcher (candidate) directly, if interested in participating in a trial that may involve the offer of acupuncture treatment *[Appendix 16]*. Those women who appeared to meet the inclusion criteria were then asked to consult their GP. In instances where the woman's GP was not participating in the trial, the researcher wrote enclosing the necessary documents for them to facilitate referral for an individual patient *[Appendix 17]*, and a follow-up phone call was made to the GP prior to their patient's planned consultation. The GP then applied the entry criteria, faxed the pro-forma to the researcher, and provided the patient with the information sheet about the study. The GP was also asked to record the diagnosis for patients who did not meet the entry criteria, and to record the reason given for patients who declined to consider further participating in the trial.

After nine months, 33 of a target of 40 patients had been recruited. This shortfall provided an opportunity to pilot a second patient-centred recruitment strategy, and in so doing maximise the chances of meeting the recruitment target within the extended 11-month recruitment period:

Patient recruitment pathway C

The GP advisor to the trial (MF) agreed to pilot a retrospective database recruitment strategy. Her staff identified patients listed on the practice computer as having consulted for treatment for idiopathic menorrhagia within the past six months and writing to invite them to contact the candidate directly. The candidate drafted this letter with the GP advisor and it was printed on the practice's headed paper *[Appendix 18]*. Enclosed with the letter were a brief information sheet *[Appendix 19]*, a reply slip and stamped envelope addressed to the candidate. The candidate then wrote to the patients who responded enclosing a copy of the patient information booklet *[Appendices 20 & 14]*.

Assignment – patient recruitment pathways A, B and C

The candidate contacted all patients recruited through pathways A, B and C and obtained informed consent from those patients who decided to enter the trial. Baseline measures were completed before the candidate contacted the remote randomisation service provided by the Department of Health Sciences, York University. This ensured unbiased allocation to trial groups and the method of random permuted blocks ensured a relatively even flow of participants to treatment groups and concealed the random allocation code. The candidate informed participants of their allocations immediately, arranged the first consultation for those randomised to and accepting the offer of acupuncture, and wrote to the practitioner to confirm this appointment, enclosing the patient's ACUMEN treatment booklet *[Appendix 21]*.

4.4 The "comparator" intervention – usual GP care

The active comparison was "usual care" in general practice, rather than "best care" according to guidelines. This option was taken in order to maximise recruitment, by not asking GPs to comply with procedures that differed from their normal practice, or that denied professional autonomy (Ross et al. 1999). How usual GP care differed from best care (Prentice 1999) was an outcome of interest.

4.5 The "experimental" intervention - Acupuncture in addition to standard GP care

4.5.1 The style of acupuncture used in ACUMEN

The ACUMEN study focused on traditional Chinese medical (TCM) acupuncture as practiced by members of the British Acupuncture Council (BAcC). BAcC members constitute the largest group of practitioners in the UK with over 2,500 members (Mills & Peacock 1997). They use traditional acupuncture as their main therapeutic intervention, have full professional indemnity insurance and work to standards of training, competency and safety set and monitored by the British Acupuncture Accreditation Board (BAAB) *[Chapter 1, 1.3]*. One clear advantage from using acupuncturists with reasonably homogenous methods of practice is that the results may be generalisable to practitioners registered with the BAcC nationally.

4.5.2 The acupuncture practitioners

Three practitioners were recruited to this study, so as to improve generalisability and avoid testing the approach of a single practitioner. The practitioners letters of formal consent were included as part of the application for Local Research Ethics Committee approval *[Appendices 22, 23, 24]* and confirmed their BAcC status, indemnity insurance, agreement to provide treatment for between five and ten patients using acupuncture as their main therapeutic intervention, and to maintain a comprehensive record of treatment in the booklets provided *[Appendix 21]*; they also agreed to participate in an interview with the candidate. All three of the practitioners had in excess of 15-years post-qualification experience. All had trained in traditional acupuncture, whilst one also had a background in Five-Elements acupuncture. They were all teachers or guest lecturers at accredited colleges and practiced at one of three licensed clinical centres in York. They were all "generalists" with a special interest either in gynaecology, paediatrics, or pain. One regularly and one occasionally used Chinese Herbal Medicine alongside acupuncture and one used kinesiology. Two of the three practitioners were female. All had previously participated in the York Acupuncture for Back Pain trial (Thomas 1999) where they had proved enthusiastic participants (Thorpe 2001).

4.5.3 The treatment rationale, regimen, needling details and co-interventions

It is generally accepted that treatment for menstrual problems is required over the course of a minimum of three menstrual cycles, involving between 12 and 20 sessions of acupuncture (Lyttleton 2004; Maciocia 1998). Patients in the experimental group were therefore offered an initial consultation and treatment and up to 19 further treatments within a six month period. Treatments were usually given once a week, certainly no more than twice a week; with the number and frequency determined by clinical need as judged by the acupuncturist. The initial consultation included a full history and a differential diagnosis was made according to medical principles originating in China. This diagnosis identified the syndrome patterns underlying each patient's symptoms of menorrhagia. The selection of points was individualised for each patient at each treatment, and often changed over time with changes in the patient's health state and response to treatment. Pre-sterilised, disposable, stainless steel needles were used according to the sterile technique defined in the British Acupuncture Council's Code of Practice. The depth of insertion varied between points, they were stimulated manually and deqi (the distinctive sensation of acupuncture) was normally sought at each point. Needles were retained for approximately 20 minutes. Other interventions such as moxibustion, acupressure-massage and cupping were administered according to theory. Similarly advice on diet, rest and exercise was based on diagnostic criteria.

4.6 Timing of follow-up and strategies to minimise loss to follow-up

Patients completed questionnaires on general health, menorrhagia, medication usage and treatment satisfaction before randomisation in the presence of the researcher, and at three, six and 12-months thereafter by post. At baseline the questionnaire also included questions about demographic characteristics, and at three, six and 12 months the questionnaires asked questions designed to monitor side effects or adverse events from treatment, satisfaction with treatment, and use and expenditure on healthcare – both private and NHS because of symptoms relating to the menstrual cycle.

The follow-up rate was an outcome of interest, as the sample size required for a full-scale trial will need to take this factor into account in order that statistical power is not lost. To maximise the follow-up rate, the front page of the questionnaire gave the trial title, identity of the organisations carrying out the research, guaranteed confidentiality, and gave the contact details

for the candidate. The questions were organised so as to avoid splitting questions and ordered logically. A consistent design was employed for similar question types and a distinct typeface with a font size of 12 points was used. Enclosed with each questionnaire was a cover letter with a picture of the candidate to aid familiarity. This letter highlighted the value of each patient's responses, outlined the nature of the questions, the predicted time required to complete, and the support available from the candidate to do this *[Appendices 25A &B; 26 & 26]*. The comparator group were also informed of the prize draw for those who completed and returned both the three and six month follow-up questionnaires *[Appendices 28A & B]*, and all patients were informed of the prize draw for those who returned their 12-month follow-up questionnaire *[Appendices 29 & 30]*. The winner of the comparator group prize draw was announced in the fourth ACUMEN newsletter that was sent to patients, acupuncturists and GPs *[Appendix 12]*. A key aim of this newsletter was to maximise response rates for the 12-month follow-up questionnaire. The candidate telephoned all non-responders after seven to ten days and offered to post them another questionnaire. After three weeks, non-responders were offered a telephone interview with the candidate to complete the measures.

4.7 Primary patient assessed outcomes

The questionnaires included the Aberdeen Menorrhagia Questionnaire and the short-form 36-item health questionnaire (SF-36). The Aberdeen Menorrhagia Questionnaire (AMQ) is a condition specific measure that is responsive to small, but clinically significant changes in the patient's condition (Ruta et al 1995). The SF-36 is a popular health profile outcome measure that has been validated for use in the NHS and gives a distinctive profile for menorrhagia (Garratt et al 1993). Thus the AMQ and SF-36 have been found to be valid, reliable and responsive to changes in patients with menorrhagia.

The Aberdeen Menorrhagia Questionnaire (AMQ)

The AMQ gives ordinal scores to all valid responses (Ruta et al 1995). The responses to each question are then summed and converted to a percentage to produce a menorrhagia severity score, between 0 and 100 *[Table 21]*. Questions refer to length of bleed, number of sanitary towels used, staining of outer clothing and average effect of symptoms on daily living over the past three months.

Table 21: The Aberdeen Menorrhagia Questionnaire - questions and scores

Ruta DA, Garratt AM, Chadha YC, Flett GM, Hall MH, Russell IT 1995 Assessment of patients with menorrhagia: how valid is a structured clinical history as a measure of health status? *Quality of Life Research* 4: 33-40

1. On average during the last 3 months, for how many days did your period last?

Less than 3 days	1
Between 3 and 7 days	2
Between 8 and 10 days	3
More than 10 days	4

2. On average during the last 3 months, were your periods regular or irregular?

Regular	1
Irregular	2

3. On average during the last 3 months, how many days were there from the first day of one period to the first day of the next period?

Less than 21 days	3
Between 21 and 35 days	2
More than 35 days	1

4. On average during the last 3 months, would you describe your periods as?

Light	1
Moderate	2
Heavy	3
Very heavy	4

5. On average during the last 3 months, for how many days of each period was the bleeding heavy?

Not heavy	1
Between 1 and 3 days	2
Between 4 and 6 days	3
Between 7 and 10 days	4
More than 10 days	5

6. On average during the last 3 months, have your periods been associated with any pain?

No pain at all	1
Slight pain	2
Moderate pain	3
Severe pain	4
Very severe pain	5

7. On average during the last 3 months, have you had any problems with soiling/staining any of the following because if your periods?

No not at all	1
Soiling/staining of outer-clothes/over-garments	2
Soiling/staining of your bed linen	3
Soiling/staining of your upholstery	4

8. On average during the last 3 months, have your periods prevented you from carrying out your work or housework or other daily activities?

No not at all	1
I could continue to work but my work suffered	2
Yes usually for no more than one day with each period	3
Yes usually for more than one day with each period	4

9. On average during the last 3 months, have you been confined to bed with each period?

No not at all	1
Yes for part of one day	2
Yes for the whole of one day	3
Yes for more than one day	4

Cont...

Table 21 continued:

10. On average during the last 3 months, have your leisure activities been affected by your heavy periods? (including social life, hobbies and sport)

Not affected by heavy periods	1
Mildly affected by heavy periods	2
Moderately affected by heavy periods	3
Severely affected by heavy periods	4
Heavy periods prevented any social life at all	5

11. On average during the last 3 months, has your sex life been affected by your heavy periods?

Not affected by heavy periods	1
Mildly affected by heavy periods	2
Moderately affected by heavy periods	3
Severely affected by heavy periods	4
Heavy periods prevented any sex life at all	5
Does not apply	6

12. On average during the last 3 months, how many tampons might you use on the heaviest day of your period?

No tampons at all	1
Between 1 and 5 tampons	2
Between 6 and 10 tampons	3
Between 11 and 15 tampons	4
More than 15 tampons	5

13. On average during the last 3 months, how many sanitary towels might you use on the heaviest day of your period?

No sanitary towels at all	1
Between 1 and 5 sanitary towels	2
Between 6 and 10 sanitary towels	3
Between 11 and 15 sanitary towels	4
More than 15 sanitary towels	5

14. At any time during the last 3 months did you require more than one form of protection at the same time (not including mini-sanitary towels or panti-liners)?

No	1
Tampon and pad together	2
Two pads together, or tampon and two pads together	3
More protection than this (i.e. disposable nappies, towels etc.)	4

Qualifications to AMQ scoring system made in ACUMEN:

1) ***Missing values given an average score***
2) ***Q.11 "does not apply" given an average score (6=3)***

Amendments to AMQ made in ACUMEN:

1) ***Q.12 is a 'Ghost question'- it was omitted from the definitive questionnaire but is included here to meet the requirements of face validity – it did not contribute to severity scores***
2) ***If Q.14 = 4, Q.13 (1 thru 3) = 4***

A patient answering with the most severe response category for each question (e.g. her periods lasted longer than 10 days with heavy bleeding, required the use of more than 15 sanitary towels on the heaviest day, caused soiling or staining of upholstery, and prevented daily activities for more than one day within each period) would score 100. Conversely, a patient ticking the least severe response category for each question (e.g. her periods lasted less than three days without heavy bleeding, had a normal cycle, required minimal protection and symptoms did not affect daily activities) would score 0. In contrast to the SF-36, the lower the score the better. Missing values were given an average score, as was the response *"does not apply"*, when asking about the impact of symptoms upon their sex life. Ruta and colleagues found scores to be positively related to GPs' perception of symptom severity, in that "severe" patients had a mean score of 50.7 (SD 21.7), "moderate" patients of 46.7 (SD 16.5), and "mild" patients 37.6 (SD 14). These GP-assessed menorrhagia severity scores were used to assess clinically significant changes.

In ACUMEN the version of the AMQ recommended by Ruta et al (1995) was adapted for the following reasons: first, for face validity, the question from the original AMQ asking women how many tampons they might use on the heaviest day of their period was retained *[Table 21, question 12]*; this question had been omitted from the definitive questionnaire as it had an item correlation of just 0.05 and did not yield a significant coefficient within one of the important factors (0.68). Thus, it was included here merely as a "ghost question", and did not contribute to women's severity scores. Second, the SPSS syntax for questions 13 and 14 was modified to reflect severity better. Several women answered *"no sanitary towels at all"* to question 13 and stated in question 14 that they used *"more protection than this, i.e. nappies"*, giving a score of just four out of a possible score of nine. Thus the syntax needs to ensure that where women answer *"more protection than this"*, they achieve a total score of at least eight or nine for questions 13 and 14. An automatic score of four, as opposed to five, was chosen for question 13, as it was unusual for women in ACUMEN to use *"more than 15 sanitary towels on the heaviest day"*.

The Short Form 36-item health questionnaire (SF-36)

The SF-36 is a comprehensive generic measure that measures health related quality of life across two summary components (physical and mental) and eight individual dimensions: (1) physical functioning; (2) role limitations due to physical problems; (3) bodily pain; (4) general

health; (5) vitality; (6) social functioning; (7) role limitations due to emotional problems; and (8) mental health (Ware et al 1993) *[see questionnaires, appendices 26, 27 & 30].* Response categories within each domain of the instrument are pre-scored, with scores ranging from zero to 100 (worst to best). SPSS syntax facilitated item aggregation and calculation of the two component and eight dimension scores. A change of between five and 10 points on the SF-36 dimension scores is widely thought to represent a clinically significant benefit (Ware et al 1992).

4.8 Additional questionnaire contents

In addition to the condition specific (AMQ) and generic (SF-36) primary patient-assessed outcome measures described above, the ACUMEN questionnaires included:

i. Questions about demographic and gynaecological characteristics,
ii. Further questions about menstrual symptoms,
iii. Two general questions that acted as global indicators of change,
iv. Questions about treatment priorities and satisfaction with care,
v. Questions devised to assess treatment acceptability and preferences,
vi. Questions about side effects or responses to treatment, and adverse events from acupuncture.

The contents of the pre-randomisation, three, six and 12-month follow-up questionnaires are described below and given in Tables 22, 23 and 24, along with the source or rationale for item inclusion. In this way the candidate sought to satisfy the requirements for face and content validity.

4.8.1 Demographic and gynaecological characteristics

The pre-randomisation questionnaire asked questions about the patient's age, ethnicity, marital status, number of children, employment, education and housing. This information was collected in order to describe the sample and gauge generalisability to other populations of women with menorrhagia. Differences in health have been found to relate to an individuals socio-economic status (Kunst & Mackenbach 1995). To provide some indication of this, women were asked information relating to their household's accommodation and their

Table 22: Contents of the ACUMEN trial participant **Pre-Randomisation questionnaire** and source / rationale for inclusion

Questionnaire / measure	Items	Source / rationale for inclusion
Demographic characteristics	1. What is your date of birth? 2. What is your age? 3. What is your marital status? Single, married, living as married, separated, divorced, widowed 4. How many children do you have? 5. What type of accommodation does your household occupy? Detached house or bungalow, semi-detached house or bungalow, terraced house (including end terrace), apartment/flat, maisonette, caravan or other mobile or temporary structure 6. Does your household own or rent the accommodation? Owns outright, owns with a mortgage or loan, pays rent and part mortgage (shared ownership), rents, lives here rent free 7. How old were you when you left full-time education e.g. school, college or university? Age 16 or less, age 17 – 19, age 20 or over, I am still in full-time education 8. Since leaving school, college or university have you had any more full or part-time further or higher education? 9. Are you? Full-time looking after the home and family and not looking for paid employment, employed full-time, employed part-time, unemployed, unable to work because of poor health 10. What is your ethnic group? Are you White: English, Irish, Scottish, Welsh, Other Are you Mixed: White and Black Caribbean, White and Black African, White and Asian, Other Are you Asian or Asian British: Indian, Bangladeshi, Other Are you Black or Black British: Caribbean, African, Other Are you Chinese or other ethnic group: Chinese, Other	1. ACUMEN trial team 2. ACUMEN trial team 3. ACUMEN trial team, CENSUS 2001 4. Shaw & Soutter 1992 5. CENSUS 2001, ACUMEN trial team 6. CENSUS 2001 7. UK BEAM 8. UK BEAM 9. UK BEAM 10. CENSUS 2001
SF-36		Brazier 1992, Garrat 1993
AMQ		Ruta 1995
Peri-Menstrual Symptoms Questionnaire (PMSQ)	Over the past 3 months how troublesome (not at all, slightly, moderately, very, extremely) were the following period symptoms? Period pain/cramps, backache, tender or swollen breasts, bloating, fluid retention, migraine or headache, tiredness or fatigue, irritability, tension, feeling low or depressed, anxiety, difficulty concentrating, crying bouts, clumsiness, aggression, other	O'Brien 1992 CATH 1 1990 Garrat 1993
York Menorrhagia Relevant Profile – YMRP	1. At your most recent period, how troublesome (not at all, slightly, moderately, very, extremely) were the following symptoms? Heavy bleeding, period pain/cramps, tiredness or fatigue 2. At your most recent period, for how many days did your period prevent you from carrying out your work, looking after your home and family, or attending college /university?	1. Garrat 1993 ; Deyo 1997 2. Deyo 1997 (Economic protocol)
Gynaecological characteristics / medication and supplements	1. For how long have you experienced heavy periods? Less than 6 months, 6 months to 1 year, 1 to 2 years, 2 to 5 years, more than 5 years 2. What factors do you think were related to the start of your heavy periods? 3. Have you ever had any of the following tests or gynaecological operations because of your periods?	1. CATH 1 2. ACUMEN trial team 3. CATH 1 4. CATH 1

used / interventions planned	No, Blood tests, D&C, hysteroscopy, laparoscopy, pelvic ultrasound, other 4. Has your GP or specialist asked you to consider having any of the following surgical treatments? No, hysterectomy, endometrial ablation, endometrial resection, hysteroscopy, laparoscopy, other 5. Are you currently on the waiting list for any of the following surgical treatments? No, hysterectomy, endometrial ablation, endometrial resection, hysteroscopy, laparoscopy, other 6. Do you use any of the following contraceptives? No, COCP, mini pill, Mirena Coil or IUS (progestogens containing coil or intra-uterine system), Depot Provera (contraceptive injection), IUCD (copper coil or intra-uterine contraceptive device) 7. Have you ever had any of these medicines for your heavy periods, and how helpful (not helpful, helped a bit, helped a lot, side effects a bit of a problem, stopped because of side effects) did you find these? (a)NSAIDS: Mefenamic acid or Ponstan, Naproxen or Naprosyn, Ibuprofen or Brufen or Nurofen (b) COCP (c) Progestogen: Norethisterone (d) Tranexamic acid: Cyklokapron (e) Progestogen containing coil: Mirenal Coil, IUS (f) Ethamsylate: Dicynene (g) Danazol (h) Other 8. What medicines did you take at your most recent period? (a) Were they prescribed by your GP, gynaecologist or self prescribed? (b) For how many days did you take this? (c)How many times per day? (d) How many tablets did you take each time? 9. Are you currently having any other treatments for your periods? 10. Do you have any other health problems apart from your heavy periods?	5. CATH 1 6. CATH 1 7. ACUMEN trial team 8. ACUMEN trial team 9. ACUMEN trial team 10. ACUMEN trial team
Global markers of change	1. In what ways have your periods changed over the last 3 months? Vastly improved, much improved, slightly improved, no change, slightly worsened, much worsened, vastly worsened 2. If you were told your periods would be like they are now until your menopause how would you feel about it? Terrible, unhappy, mostly dissatisfied, Mixed, Mostly satisfied, pleased, delighted	Deyo 1997
Treatment priorities	1. How important is reducing your menstrual blood loss? (very important, fairly important, slightly important, not important) 2. How important is increasing your energy levels? 3. How important is increasing the activities you can do? 4. How important is cutting down or avoiding medication? 5. How important is avoiding surgery?	1. Ruta 1995 2. Garrat 1993 3. Ruta 1995 4. MYMOP 5. Marshall 1998, Longridge nee Gamon 2000
Satisfaction with treatment and care (1)	1. Over the past 3 months, how satisfied (very satisfied, somewhat satisfied, neither satisfied nor dissatisfied, somewhat dissatisfied, very dissatisfied) were you with the following aspects of your treatment and care? The information you were given about your heavy periods The treatment you received for your heavy periods The overall care you received for your heavy periods	ACUMEN trial team
Satisfaction with treatment and care (2)	1. Over the past 3 months, how would you rate the following (poor, fair, good, very good, excellent)? The explanation of your condition and how treatment might help? The explanation given about what treatment would involve for you? The attention given to what you have to say? The advice given to you about ways to become and stay healthy? The friendliness and courtesy shown to you?	Cassidy 1998b Luff & Thomas 2000 Gould & MacPherson 2001 YACBAC

	The respect shown to you, such as attention to your privacy? The reassurance and support offered to you? The time give to you at each visit for advice and/or treatment?	
Previous use of CAM & Beliefs concerning acupuncture	1. Have you ever consulted a complementary or alternative medicine practitioner, for example, one of those listed in the box below? If yes, which practitioner did you consult (acupuncturist, chiropractor, homeopath, herbalist, osteopath, other)? If yes, was it for treatment for your heavy periods or another health condition? 2. In general, do you think acupuncture can work? (Yes, no, don't know) 3. Do you think your heavy periods might be helped by acupuncture (Yes, no, don't know) 4. If you were offered acupuncture treatment for your heavy periods would you choose to accept or decline this offer? (accept, decline, don't know)	Torgerson 1996 House of Lords 2000 Wadlow & Peringer 1996

References for table 4

Brazier et al 1992 - Brazier JE, Jones N, Kind P 1992 Testing the validity of the EuroQol and comparing it with the SF-36 health survey questionnaire *Quality of Life Research* 2: 169-180

CATH 1 - Pinion S 2001 Personal communication regarding the Conservative Alternatives to Hysterectomy (CATH 1) original trial questionnaire

Cassidy 1998b – Cassidy CM. Chinese medicine users in the United States part 2: preferred aspects of care. *The Journal of Alternative and Complementary Medicine* 1998; 4 (2): 189-202

CENSUS 2001 *29 April "count me in" census 2001: England household form – Specimen* www.statisitics.gov.uk/census

Coulter 1995 - Coulter A, Kelland J, Peto V, Rees MCP 1995 Treating menorrhagia in primary care: an overview of drug trials and a survey of prescribing practice *International Journal of Technology Assessment in Health Care* 11; 3: 456-471

Deyo 1997 – Deyo R, Battie M, Beurskens AJHM, Bombardier C, Croft P, Koes B, Malmivaara A, Roland M, Korff MV, Waddell G 1998 Outcome measures for low back pain research: a proposal for standardised use. *Spine* 23; 18: 2003-2013

EuroQol Group 1991 - EuroQol: a new facility for the measurement of health related quality of life *Health Policy*

Garrat et al 1993 - Garratt AM, Ruta DA, Abdalla MI, Buckingham JK, Russell IT 1993 The SF-36 health survey questionnaire: an outcome measure suitable for routine use within the NHS? *British Medical Journal* 306:1440-1444

Gould & MacPherson 2001 – Gould A & MacPherson H. Patient perspectives on outcome after treatment with acupuncture. *The Journal of Alternative and Complementary Medicine* 2001; 7(3): 261-8

House of Lords 2000 - House of Lords Select Committee on Science and Technology 2000 Complementary and Alternative Medicine Version 6 (HL Paper 123) London the Stationary Office

Longridge nee Gamon 2000. *Primary Dysmenorrhoea: exploring the potential of acupuncture.* MSc Thesis, Northern College of Acupuncture, York accredited by University of Wales

Luff & Thomas 2000 – Luff D & Thomas KT. "Getting somewhere", "feeling cared for": patients perspectives on complementary therapies in the NHS. *Complementary Therapies in Medicine* 2000; 8 (4): 253-259

MYMOP - Patterson C, Britten N 2000 In pursuit of patient-centred outcomes: a qualitative evaluation of the 'Measure yourself medical outcome profile'. *Journal of Health Services Research Policy*, 5 (1): 27-36

Marshall 1998 – Marshall J 1998 An exploration of women's concerns about heavy menstrual blood loss and their expectations regarding treatment *Journal of Reproductive and Infant Psychology* 16: 259-276

O'Brien 1992 – O'Brien P 1992 Premenstrual Syndrome In *Gynaecology* Eds. Shaw R, Soutter WP Churchill Livingstone

Ruta et al 1995 - Ruta DA, Garratt AM, Chadha YC, Flett GM, Hall MH, Russell IT 1995 Assessment of patients with menorrhagia: how valid is a structured clinical history as a measure of health status? *Quality of Life Research* 4: 33-40

Shaw R, Soutter WP Eds. 1992 *Gynaecology* Churchill Livingstone

Torgerson 1996 - Torgerson DJ, Klaber-Moffett J, Russell IT. Patient preferences in randomized trials: threat or opportunity? *Journal of Health Services Research Policy*, 1996; 1 (4): 194-197

UK BEAM trial participant randomisation questionnaire 2000 – Department of Health Sciences, University of York

Wadlow & Peringer 1996 – Wadlow G, Peringer E 1996 Retrospective survey of patients and practitioners of traditional Chinese acupuncture in the UK *Complementary Therapies in Medicine* 4: 1-7

YACBAC trial participant questionnaire 1999 – Medical Care Research Unit, University of Sheffield and Foundation for Traditional Chinese Medicine, York

education, because accommodation is a shared attribute and so gives a sense of the family's shared status, whilst education is a personal attribute and so provides a better indicator of the individual's socio-economic status. In particular, it asked each woman's age and whether or not she had had children, as these are known prognostic variables: adolescent girls and women over 40-years are at greater risk of anovulatory menorrhagia, which gives the heaviest bleeds, and parous women tend to have heavier less painful menses than non-parous women (Shaw & Soutter 1992). These questions in the pre-randomisation questionnaire were developed with reference to a trial questionnaire (UKBEAM) and the 2001 Census questionnaire; a strategy that aimed to improve the reliability of demographic data and enable comparisons to be made to the population as a whole *[Table 22, appendix 26]*.

The pre-randomisation questionnaire also asked women whether they had any health related problems apart from menorrhagia, as concurrent conditions are confounding factors; for how long they had had heavy periods, as chronicity is a prognostic variable; and for the factors they thought were related to the onset of symptoms: The rationale being that in cases where their periods have not always been heavy, women often describe the onset of symptoms in relation to a significant life event, such as childbirth, and this is information that is relevant to the understanding of their condition from the perspective of traditional acupuncture.

In order to understand where on the treatment trajectory women joining ACUMEN were, patients were asked about 1) the diagnostic tests they had received, including blood tests, laparoscopy and pelvic ultrasound; 2) whether or not their GP or a specialist had asked them to consider a surgical investigation such as hysteroscopy, or a surgical option such as hysterectomy and endometrial ablation; 3) and what prescribed medicines they had used to date (e.g. oral contraceptive pill and norethisterone). To understand treatment usage at baseline, women were asked what contraceptives they were currently using, as these can both reduce blood loss and pain (e.g. the contraceptive pill and MIRENA coil) and increase blood loss (e.g. the intra-uterine contraceptive device), as well as the medications, supplements and therapies they had used at their most recent period (e.g. iron supplements, paracetamol, tranexamic acid and homeopathy) *[Table 22, appendix 26]*. These questions were informed by those in a menorrhagia trial questionnaire (CATH 1), with similar questions asked at follow-up in order to monitor patients' treatment usage and their referral to secondary care for surgical investigations and treatment *[Table 23, appendices 27 & 30]*.

Table 23: Contents of the ACUMEN trial **3-month follow-up questionnaire** and rationale for item inclusion

Questionnaire / measure	Items	Source / rationale for inclusion
SF-36		
AMQ		
PMSQ	*See table regarding pre-randomisation questionnaire contents – Peri-Menstrual Symptoms questionnaire*	
YMRP	*See table regarding pre-randomisation questionnaire contents – York Menorrhagia Relevant Profile*	
Treatments used, investigations and surgery planned	1. Since joining the study 3 months ago, have you had any of the following tests or gynaecological operations because of your periods? - No, Blood tests, D&C, hysteroscopy, laparoscopy, pelvic ultrasound, other 2. Since joining the study 3 months ago, has your GP or specialist asked you to consider having any of the following surgical treatments? - No, hysterectomy, endometrial ablation, endometrial resection, hysteroscopy, laparoscopy, other 3. Since joining the study 3 months ago, have you been put on the waiting list for any of the following surgical treatments? No, hysterectomy, endometrial ablation, endometrial resection, hysteroscopy, laparoscopy, other 4. Since joining the study 3 months ago, has the type of contraceptive you used changed? 5. If yes, do you now use any of the following contraceptives? No, COCP, mini pill, Mirena Coil or IUS (progestogens containing coil or intra-uterine system), Depot Provera (contraceptive injection), IUCD (copper coil or intra-uterine contraceptive device) 6. What medicines did you take at your most recent period? (a) Were they prescribed by your GP, gynaecologist or self prescribed? (b) For how many days did you take this? (c)How many times per day? (d) How many tablets did you take each time? 7. Have you had any other treatments for your periods – not including study acupuncture?	1. CATH 1 2. CATH 1 3. CATH 1 4. CATH 1 5. CATH 1 6. ACUMEN trial team 7. ACUMEN trial team *These are follow-up versions of 7 of the 10 questions asked prior to randomisation*
Global markers of change	*See table 2 – pre-randomisation questionnaire contents*	
Treatment side effects, responses or sensations	1. Over the past 3 months, during which time you have received treatment for your heavy periods, have you experienced any of the following? (Yes, no) – Feeling drowsy, feeling dizzy, vertigo, feeling faint, nausea, vomiting, irregular periods, erratic periods, absence of periods, depression, headache, skin problems, diarrhoea, stomach pain, breast tenderness, water retention, other?	1. BNF 2000 2. MacPherson 2000 (2001)
Adverse events	*1. Over the past 3 months have you received acupuncture for your heavy periods? As part of this research project? Or Private acupuncture? – If yes, have you experienced (a) pain where the needle was inserted (b) bruising where the needle was inserted (c) bleeding where the acupuncture needle was inserted for more than 10 seconds (d) other...*	1. MacPherson 2000 (2001)
Priorities	*See table regarding pre-randomisation questionnaire contents – for questions concerning treatment priorities*	
Care (1)	*See table regarding pre-randomisation questionnaire contents – for questions about satisfaction with treatment and care (1)*	
Care (2)	*See table regarding pre-randomisation questionnaire contents– for questions about satisfaction with treatment and care (2)*	

References for table 23 (see also table 22)

BNF 2000 - *British National Formulary* (March) 2000 British Medical Association and Royal Pharmaceutical Society of GB, London

CATH 1 - Pinion S 2001 Personal communication regarding the Conservative Alternatives to Hysterectomy (CATH 1) original trial questionnaire

MacPherson 2000 (2001) – Personal communication regarding study published in 2001 - MacPherson H, Thomas K, Walters S, Fitter M 2001 The York acupuncture safety study: prospective survey of 34,000 treatments by traditional acupuncturists *British Medical Journal* 323: 486-487

4.8.2 Further questions about menstrual symptoms

Garrat and colleagues (1993) found that the dimensions of the SF-36 most adversely affected by menorrhagia are pain and vitality. The AMQ (Ruta et al 1995) measures menorrhagia severity in terms of blood loss, pain, and impact upon daily living over the previous three months, but does not ask about fatigue or vitality directly. The York Menorrhagia-Relevant Profile (YMRP) was devised in order to focus upon and measure the three symptoms of specific relevance to menorrhagia – heavy bleeding, pain and fatigue. It was informed by the questionnaire for low back pain developed by Deyo and colleagues (1992), whereby it complemented questions that had looked at a longer time period by focusing instead on the most recent time period. For Deyo's questionnaire this had meant asking about the *"troublesomeness"* of specific back pain symptoms on a five-point scale (i.e.*"extremely" "very" "moderately" "slightly"* and *"not at all* troublesome") over the previous week, as compared to the past month. In this study this had meant the last menstrual cycle, as compared to the previous three months asked in the AMQ.

This five-point scale question format has been found to give reliable results (Streiner & Norman 1995; Deyo 1997). Thus the Peri-Menstrual Symptoms Questionnaire (PMSQ) used this same question format and asked about 15 common symptoms including pain, fatigue, breast tenderness and irritability. The items included were chosen on the basis of their inclusion in a list of common premenstrual symptoms in a consensus medical text (O'Brien in Shaw & Soutter eds. 1992) and their inclusion in a previous menorrhagia study questionnaire - CATH 1 (Pinion 2001).

The YMRP and PMSQ were devised to estimate the wider effects of treatment upon menstrual health. The rationale being twofold: First, such concurrent symptoms as fatigue, breast tenderness and irritability have important implications for quality of life. Second, successful acupuncture treatment would be expected to bring about improvements in the full array of menstrual symptoms experienced, not just the bleed (Maciocia 1997; Flaws 1997; Kaptchuk 1983) *[Table 22, appendices 26, 27 & 30].*

Table 24: Contents of the ACUMEN trial **six and 12-month follow-up questionnaires** and rationale for item inclusion

Questionnaire / measure	Items	Source / rationale for inclusion
SF-36		
AMQ		
PMSQ	*See table re: pre-randomisation questionnaire contents – for Peri-Menstrual Symptoms questionnaire*	
YMRP	*See table re: pre-randomisation questionnaire contents- for Your Most Recent Period questionnaire*	
Treatments used etc.	See table re: 3-month follow-up questionnaire contents – for questions concerning treatments used and investigations planned etc	
Global markers of change	*See table re: pre-randomisation questionnaire contents*	
Treatment side effects, responses or sensations	*See table re: pre-randomisation questionnaire contents – for questions concerning side effects, responses or sensations from treatment*	
Adverse events	*See table re: pre-randomisation questionnaire contents – for questions concerning adverse events from acupuncture*	
Priorities	*See table re: pre-randomisation questionnaire contents – for questions concerning treatment priorities*	
Care (1)	*See table re: pre-randomisation questionnaire contents – for questions about satisfaction with treatment and care (1)*	
Care (2)	*See table re: pre-randomisation questionnaire contents – for questions about satisfaction with treatment and care (2)*	
Treatment acceptability and preferences	How strongly do you agree or disagree with the following statements – (strongly agree, agree, uncertain, disagree, strongly disagree) - 1. Knowing what I know now about treatments for heavy periods, if I could travel back in time 6 / 12 months I would want to have (a) acupuncture (b) prescribed medication (c) hysterectomy (d) other 2. Knowing what I know now about treatments for heavy periods, if in the future (tomorrow or in two years time) I needed treatment for my heavy periods, I would want to have (a) acupuncture (b) prescribed medication (c) hysterectomy (d) other 3. Knowing what I know now about treatments for heavy periods, I would recommend to other women (a) acupuncture (b) prescribed medication (c) hysterectomy (d) other	ACUMEN trial team
Health care costs	*Included in appendix 28*	

References for table 14 (see also table 22)

BNF 2000 - *British National Formulary* (March) 2000 British Medical Association and Royal Pharmaceutical Society of GB, London

CATH 1 - Pinion S 2001 Personal communication regarding the Conservative Alternatives to Hysterectomy (CATH 1) original trial questionnaire

MacPherson 2001 - MacPherson H, Thomas K, Walters S, Fitter M 2001 The York acupuncture safety study: prospective survey of 34,000 treatments by traditional acupuncturists *British Medical Journal* 323: 486-487

4.8.3 Global markers of change

The first of these two questions asked *"in what ways have your periods changed over the past three/six months?"* and aimed to understand the direction of change in symptoms women experienced at baseline and at each follow-up *("vastly improved", "much improved", "slightly improved", "no change", "slightly worsened", "much worsened", "vastly worsened").* The second question asked *"If you were told your periods would be like they are now until your menopause how would you feel about it? – "terrible", "unhappy", "mostly dissatisfied", "mixed", "mostly satisfied", "pleased", "delighted".* The aim was to understand the extent to which women's treatment needs had been met. This question was adapted from a question in the Deyo back pain questionnaire (1997) – *"if you had to spend the rest of your life with your back condition as it is now, how would you feel about it?"* Both of the questions used a seven point scale, so enhancing the reliability of the findings (Streiner & Norman 1995) *[Table 22, appendices 26, 27 & 30].*

4.8.4 Treatment priorities and satisfaction with care

The first of this group of questions asked women to rate on a four point scale *("not important", "slightly important", "fairly important", "very important")* the importance of reducing menstrual blood loss, increasing energy levels, increasing activities, cutting down or avoiding medication, and avoiding surgery. These questions aimed to understand women's treatment priorities and were informed from the literature: Ruta and colleagues (1995) had found excessive menstrual loss to be an important symptom that has a profound impact upon the activities women engage in; Garrat and colleagues (1993) described the impact of symptoms upon the SF-36 dimension "vitality"; Patterson and Britten (2000) had found that people often choose a CAM modality in order to reduce their medication usage or avoid medication, whilst Marshall (1998) and Longridge nee Gamon (2000) had found that some women would like to avoid both medication and surgery *[Table 22, appendices 26, 27 & 30].*

Women were also asked to rate on a five point scale *("very satisfied", "somewhat satisfied", "neither satisfied nor dissatisfied", "somewhat dissatisfied" "very dissatisfied")* their satisfaction with aspects of their treatment and care: the information they were given about their heavy periods; the treatment they received for their heavy periods; and how they would rate the overall care they received. A more detailed questionnaire then asked women to rate on

a five point scale *("poor", "fair", "good", "very good", "excellent")* their experience of different elements of the treatment encounter: the explanation they'd been given of their condition and treatment; the attention given to what they have to say; the advice given about ways to become and stay healthy; the friendliness and courtesy shown; the respect shown, such as attention to privacy; the reassurance and support offered; and time given at each treatment session. The rationale for their inclusion being that these are all factors CAM patients have sighted as important in their satisfaction with treatment and care (Luff & Thomas 2000; Gould & MacPherson 2001; Cassidy 1998b). This questionnaire was developed with reference to one used in the York Acupuncture for Back Pain (YACBAC) trial questionnaire *[Table 22, appendices 26, 27 & 30]*.

4.8.5 Treatment acceptability and preferences

To understand patients' previous use of CAM modalities (House of Lords 2000; Wadlow & Peringer 1996) and whether patients in ACUMEN held a preference for, or were indifferent to receiving either the experimental or control intervention (Torgerson et al 1996), a number of questions were asked in the pre-randomisation questionnaire. First, patients were asked if they had ever consulted a CAM practitioner, and if so whether they had sought treatment for their heavy periods or another health condition. It also asked if they thought acupuncture could work in general and if they thought it could work for heavy periods in particular. Lastly it asked patients what they would do if randomised to the offer of acupuncture, did they think they would accept or decline this offer, or were they unsure what they would choose to do *[Table 22, appendix 26]*?

At six and 12-month follow-up three questions were added in order to understand patients' views of treatment acceptability given their experiences over the course of ACUMEN. The questions were developed by the ACUMEN trial team and used a format from the SF-36 where patients are asked to report on a five point scale how strongly they agree or disagree with a statement: *"strongly agree", "agree", "uncertain", "disagree", "strongly disagree"*. The first set of three statements read *"knowing what I know now about treatments for heavy periods, if I could travel back in time six/twelve months I would want to have a) acupuncture, b) prescribed medication, c) hysterectomy, and d) other therapy"*. The second set of statements read *"knowing what I know now about treatments for heavy periods, if in the future*

(tomorrow or in two years time) I needed treatment for my heavy periods, I would want to have a) acupuncture, b) prescribed medication, c) hysterectomy, and d) other therapy". The final set of statements read *"knowing what I know now about treatments for heavy periods, I would recommend to other women a) acupuncture, b) prescribed medication, c) hysterectomy and d) other therapy".* In this way this questionnaire aimed to determine the strength of preference for a particular intervention amongst trial participants. In particular, it sought to understand whether with the benefit of hindsight those women who were offered acupuncture considered GP referral for treatment appropriate, whether they'd consider it appropriate for other women to receive this offer, and whether they would prefer to use acupuncture for their heavy periods in the future *[Table 24, appendix 30].*

4.8.6 Side effects or responses to treatment and adverse events from acupuncture

To detect and monitor side effects from or responses to treatment, a questionnaire was developed with reference first, to those side effects common to medication prescribed for menorrhagia and listed in the British National Formulary, for example, stomach pain and erratic menses (March 2000); and second with reference to those treatment responses included in the questionnaire for a prospective acupuncture safety survey, for example, feeling drowsy and headaches (MacPherson 2000). In order to detect and monitor adverse events from acupuncture needle insertion, and again with reference to the acupuncture safety survey, patients were asked whether they had experienced pain where the needle was inserted, bruising where the needle was inserted, or bleeding where the acupuncture needle was inserted for more than ten seconds *[Table 23, appendices 27 & 30].*

4.9 Sample size

The sample size of 40 patients was chosen for pragmatic reasons. First, 20 observations in each arm of the trial was expected to be sufficient to compute the sample size that would be necessary to give an 80% chance of detecting an important change in scores at a 5% significance level on the Aberdeen Menorrhagia Questionnaire (AMQ) and appropriate sub-scales of the SF-36. Second, as past research has found that in any year five per cent of all women of reproductive age will consult their general practitioner with the complaint of menstrual dysfunction (Vessey et al 1992), the recruitment of 40 patients was considered

feasible within one Primary Care Trust over five months. Third, funding was obtained for the treatment of 20 patients with acupuncture.

4.10 Masking

As ACUMEN was a pragmatic exploratory trial designed to answer questions about the relative clinical effectiveness of different treatment options for menorrhagia, including acupuncture, in the real world setting of primary care, blinding of participants and professionals was neither desirable nor possible. Blinding the evaluator was not possible because of the restraints and requirements of a PhD.

4.11 Analysis

The candidate entered all data into a statistical package (SPSS), and although recruitment and follow-up was monitored closely, the full data set was analysed only once, to avoid multiple significance testing (Campbell & Machin 1999). At baseline, categorical data were tabulated by group, whilst the quantitative data were analysed by ANOVA to estimate means and standard deviations, and to test whether differences were significant between groups. The principle comparison was between all patients randomised to receive the option of referral to the acupuncture service versus all patients randomised to usual GP management alone. The analysis therefore estimated the benefits of offering traditional acupuncture to this group of clinically defined patients. All analyses were undertaken on an intention-to-treat basis. As the distribution approximated normal in this relatively small sample the preferred option of analysis of covariance (ANCOVA) was used to control for baseline and detect treatment effect (Vickers & Altman 2001; Vickers 2005). Area-under-curve analysis was carried out to summarise findings and give most weight to patients' experiences over 12-months (Matthews et al 1990). The Sign Test was used to test for significance in the direction of change in primary and secondary outcomes at follow-up (Armitage & Berry 1987).

4.12 Research Governance

4.12.1 Details of the ACUMEN research advisory group

This exploratory trial of acupuncture for menorrhagia was a collaborative PhD project between the Department of Health Sciences (DHSc), University of York, the York-based research Foundation for Traditional Chinese Medicine (FTCM), and the Medical Care Research Unit (MCRU), University of Sheffield. Both the DHSc and MCRU have a national reputation for conducting health services research, including randomised clinical trials and qualitative research inquiries. And by the start of ACUMEN, the MCRU and FTCM had designed and implemented a pragmatic trial of acupuncture for low back pain (Thomas et al 1999). The candidate worked under the guidance and tuition of her multi-disciplinary research advisory group (RAG), who offered expertise in health technology assessment, health economics, acupuncture research, and general practice. The candidate also benefited from the input, support and commitment of a lay advisor and two further acupuncture practitioners with whom the FTCM had established links.

Prof. Ian Russell, Academic Supervisor and Professor of Health Sciences, University of York

Dr. Hugh MacPherson, Clinical Supervisor, ACUMEN acupuncturist and Research Director, Foundation for Traditional Chinese Medicine, York

Ms Kate Thomas, Academic Co-supervisor and Deputy Director, Medical Care Research Unit, University of Sheffield

Prof. Christine Godfrey, Economic Advisor and Professor in Health Economics, University of York

Dr. Manuela Fontebasso, GP Advisor, York

Ms Fenella Jeffers, Lay advisor, Foundation for Traditional Chinese Medicine, York

Ms Alison Gould, The Northern Clinic of Acupuncture, 61 Micklegate, York

Ms June Tranmer, The Healing Clinic, Fulford Cross, York

Dr Daphne Russell, Statistical Advisor and Lecturer in Health Sciences, University of York

4.12.2 Funding

The costs of the research were met by a research studentship from the University of York and a research grant from the Foundation for Traditional Chinese Medicine, which provided office premises, met the administrative costs of the research, and funded the acupuncture treatments for trial patients. Thus the ACUMEN exploratory trial did not impose treatment or service costs on the NHS.

4.12.3 York Research Ethics Committee (YREC) approval

The candidate, with the support of her supervisory team, first applied for ethical approval for the ACUMEN trial and the nested qualitative study in January 2001 *[Appendix 31]*. Approval was granted in March 2001, once the research team had successfully addressed the issues raised by the committee in response to the initial application *[Appendices 32 to 34]*. In January 2002 the candidate and supervisory team decided to pilot recruitment pathway C needing the GP advisor to write to eligible patients on her practice list *[Subsection 4.2.2]*. The candidate therefore wrote directly to the Chair of YREC with the necessary amendments to the approved protocol and received permission in March 2002 *[Appendices 35 & 36]*.

4.12.4 Informed consent

All patients invited to consider participating in the trial were provided with an information sheet by either their GP – pathways A and B – or the researcher – pathway C *[Figure 5 & Appendix 14]*. This explained the purpose of the study, the nature of the interventions – experimental and comparator, when and how the questionnaires would be administered, and the data management procedures in place to guarantee confidentiality and anonymity. It also explained that participants could make a formal complaint by contacting the candidate or by following either the NHS or British Acupuncture Council's complaints procedure. Patients following recruitment pathways A and B gave their written consent for their GP to pass on their contact details to the candidate *[Appendix 15]*. And patients following recruitment pathways A, B and C gave their written consent to the candidate if they wished to participate in the trial *[Appendix 37]*.

4.12.5 Data management

The collection and management of data was carried out in line with the Data Protection Act 1998, and in accordance with the Policy for the Protection of Data and Data Providers prescribed by the Department of Health Sciences, University of York. The participants' contact details were entered into the master index at the research centre by the candidate and protected by password. All other data held either on computer, or in a locked filling cabinet at the research centre, were coded only with the interviewee's unique study number. These data included the questionnaires and copies of the patient treatment booklets. The file containing participants' contact details will be erased after completion of the research.

5 Results

The principal findings from the ACUMEN exploratory trial are described below and in Tables 25 to 35.

5.1 Participant flow and follow-up

Thirty-nine of a target of 40 women were recruited over a period of 11-months by 43 GPs in the Selby and York Primary Care Trust – June 2001 to April 2002 - giving a patient recruitment rate of 3.5 patients per month. Twenty were randomised to the acupuncture group and 19 the usual care group. Of these, approximately 40% came in response to GP surgery posters and GP invitation (recruitment pathway A); almost half came to the study in response to other patient-centred recruitment strategies by first contacting the candidate (recruitment pathway B); and 10% came in response to letters from their GP (recruitment pathway C) *[Table 25 and Chapter Five]*. All twenty of the women randomised to acupuncture accepted this offer and completed a course of up-to 20 sessions of acupuncture within a six month treatment period. At three months, 37 (95%) returned questionnaires; at six months, 36 (92%); and at 12 months 32 (82%). At both the three and six-month follow-up, all non-responders were in the comparator treatment group, but at 12 months two participants were lost from the acupuncture group. Thus ACUMEN achieved a loss to follow up rate of 5% at three months, 8% at six months and 18% at 12 months.

Table 25: Recruitment pathways – values are numbers (percentages)

Recruitment pathway	Acupuncture Group	Standard Care Group	Total
GP surgery poster / leaflets	9 (45.0)	3 (15.8)	12 (30.8)
GP initiated within consultation	2 (10.0)	2 (10.5)	4 (10.3)
Total for pathway A	**11 (55.0)**	**5 (26.3)**	**16 (41.0)**
Local newspaper article about study	3 (15.0)	6 (31.6)	9 (23.1)
Acupuncturist initiated	4 (20.0)	2 (10.5)	6 (15.4)
ACUMEN patient recommendation		2 (10.5)	2 (5.1)
Acupuncture clinic poster / leaflets		1 (5.3)	1 (2.6)
Local radio feature about acupuncture where study was mentioned		1 (5.3)	1 (2.6)
Total for pathway B	**7 (35.0)**	**12 (63.2)**	**19 (48.7)**
Letters to a GPs patients (n=14)	2 (10.0)	2 (10.5)	4 10.3)
Total for pathway C	**2 (10.0)**	**2 (10.5)**	**4 (10.3)**
	20 (100)	19 (100)	39 (100)

5.2 Baseline data

5.2.1 Demographic characteristics

The women participating in ACUMEN were aged between 21 and 52 years, with an average age of 41 years (SD 7.13). The great majority (95%) described themselves as *"white English"*. Ninety-two percent (33/39) had had a child, and most had had two or more (72%). The majority lived in a house or bungalow, as opposed to a flat or maisonette (13%). Approximately half owned their homes with a mortgage, almost a quarter owned their homes outright (23%), and another 23% paid rent. About half had left full time education when aged less than 16 years, and most (77%) had received further education since leaving school, college or university. The majority were employed part-time (46%), whilst around 40% were either employed full-time or were full-time looking after the home and family. Three (7.7%) of the women participating were unable to work because of poor health and described a complex medical history with cerebella ataxia, depression or pleurisy being the primary health problem. Both randomised groups shared similar characteristics.

Forty-nine percent (19/39) of the women participating in ACUMEN had previously consulted a complementary or alternative medicine (CAM) practitioner, with 36% (14/39) having previously used acupuncture: five for treatment for heavy periods and nine for other health

conditions *[Table 26]*. Prior to randomisation, 69% (27/39) of participants reported that in general, they thought acupuncture could work, and 54% (21/39) thought it could work for heavy periods. None of the women reported they did not think it could work, responding instead *"don't know"*, and just one woman said she would decline acupuncture if offered – although she did in fact accept and complete a course of acupuncture as part of ACUMEN. Again, both randomised groups shared similar characteristics.

Table 26: Proportion (number) of patients previously consulting CAM practitioners

Acupunctur't	*Homeopath*	*Chiropractor*	*Osteopath*	*Herbalist*	*Kinesiologist*	*Aromathera't*
36% (14/39)	23% (9/39)	13% (5/39)	10% (4/39)	3% (1/39)	3% (1/39)	3% (1/39)

5.2.2 Gynaecological characteristics

The majority (62%; 24/39) of women in ACUMEN had experienced heavy periods for more than five years, 23% (9/39) for between two to five years, 8% (3/39) for one to two years, 3% (1/39) for between six months and a year, and 5% (2/39) had experienced symptoms for less than six months. At randomisation, none of the participants were on the waiting list for gynaecological investigations or surgery, although four (10%) had been asked to consider hysterectomy by their GP, and one (3%) to consider endometrial resection. Only one third (33%; 13/39) had used one or more prescribed drug therapies to reduce excessive menstrual blood loss at their most recent period. Eight women had not used anything – not prescribed medications, over the counter NSAID, supplements or CAM therapies. Thirty-five of the women (90%) were not using a prescribed contraceptive, one (3%) was taking a combined oral contraceptive pill, one (3%) was using the Mirena coil, and two the copper coil or intra-uterine contraceptive device (5%) *[Table 27]*. At their most recent period, 31 (80%) had taken a prescribed drug therapy or nutritional supplement and eight (21%) had not used drug therapies or supplements; twenty of the women (52%) had used an over-the-counter non-steroidal anti-inflammatory drug (NSAID); five had used tranexamic acid (13%), five mefenamic acid (13%), three norethisterone (8%), one the oral contraceptive pill (OCP) (3%), and one hormone replacement therapy (HRT) (3%). A fifth of participants (21%) had taken an iron nutritional supplement and a quarter had taken other nutritional supplements – including multivitamin preparations, Oil of Evening Primrose and Vitamin B6. Chi-squared tests found no significant difference between groups in treatments used *[Table 27]*.

Table 27: Drug therapies, nutritional supplements and CAM therapies used to address menstrual symptoms

Reported at:	Baseline		Three months		Six months		12 months	
Treatment	Acupuncture + usual care (max n=20/20)	Usual care (max n=19/19)	Acupuncture + usual care (max n=20/20)	Usual care (max n= 17/19)	Acupuncture + usual care (max n=20/20)	Usual care (max n=16/19)	Acupuncture + usual care (max n=18/20)	Usual care (max n=14/19)
OCP	0	1	0	2	0	0	0	2
MIRENA coil	0	1	0	1	1	1	2	0
Tranexamic Acid	2	3	0	2	0	2	1	1
Mefenamic Acid	2	3	1	2	1	2	0	2
Norethisterone	0	3	1	2	1	1	0	1
Thiazide diuretic	0	0	0	0	0	1	0	0
HRT	1	0	0	0	0	0	0	0
IUCD	0	2	0	2	0	2	0	2
Over-the-counter NSAID	10	10	10	11	8	6	10	10
Iron nutritional supplement	4	4	3	2	3	2	3	1
Other nutritional supplement	4	6	1	8	2	8	2	4
Homeopathy	0	1	1	1	0	1	0	0
Nutrition	0	0	1	0	0	0	0	0
ACUMEN acupuncture	0	0	20	0	20	0	0	0
Private acupuncture	1	1	0	0	0	0	3	1

Notes

a Apart from non-responders (Section 5.1), there were no missing data.

b Chi-squared tests found no significant difference between groups in treatments used.

5.2.3 *Mean baseline scores on the primary outcomes - AMQ, SF-36 and EQ-5D*

At baseline, the scores on the AMQ ranged from 25 to 78 out of 100, with a mean of 52 *[Table 28]*. Mean baseline scores on the SF-36 physical component were 44 and for the mental component 41. The mean scores for the individual dimensions were around 60 and above for all but pain (44) and vitality (39). And the mean score for the EQ-5D was 0.749. Though there were differences between groups, t tests found no significant differences between groups. (Nevertheless analysis of covariance remains the best way of taking these differences into account when analysing changes over 12 months).

Table 28: Baseline scores for primary patient assessed outcomes - [values are mean (SD) scores unless stated otherwise]

Outcome measure	**Acupuncture plus usual care** (max n=20/20)	**Usual care** (max n=19/19)	**Total sample** (max n=39/39)
AMQ (0-100, 0=best)	53.9 (12.0)	50.3 (12.5)	52.1 (12.2)
SF-36 (mean=50, SD=10, 100=best)			
Physical component score	44.9 (8.2)	44.0 (9.3) (n=18/19)	44.5 (8.6) (n=38/39)
Mental component score	41.2 (10.4)	41.2 (11.9) (n=18/19)	41.2(11.0) (n=38/39)
1. Physical functioning	79.7 (24.0)	83.2 (15.2)	82.0 (20.0)
2. Social functioning	63.1 (24.2)	65.1 (20.2)	64.1 (22.1)
3. Role physical	66.9 (28.8)	61.8 (27.6)	64..4 (28.0)
4. Role mental	73.3 (25.3)	76.8 (20.7)	75.0 (22.9)
5. Mental health	61.7 (19.9)	59.7 (19.0) (n=18/19)	60.8 (19.2)
6. Vitality	40.3 (18.7)	37.5 (21.0) (n=18/19)	39.0 (19.6) (n=38/39)
7. Pain	44.4 (27.2)	42.6 (21.9)	43.6 (24.5) (n=38/39)
8. General health	58.8 (18.0)	56.3 (15.0) (n=18/19)	57.6 (16.5)

Notes
t tests found no significant differences between groups

5.2.4 *Mean baseline scores on the secondary outcomes - YMRP and PMSQ*

At baseline, women scored the *"troublesomeness"* of three key symptoms on a five point scale (5=worst) at their *most recent* period *[Table 29].* The mean score for the dimension

Table 29: Baseline scores on secondary patient assessed outcomes the YMRP and PMSQ – [values are mean (SD) scores unless stated otherwise]

Outcome measure	**Ac. plus standard care** (max n=20/39)	**Standard care** (max n=19/39)	**Total** (max n=39/39)
YMRP (1-5, 1=best)			
1. Heavy bleeding	3.9 (0.9)	3.4 (1.0)	3.6 (1.0)
2. Period pain	3.6 (1.1)	3.2 (1.0)	3.4 (1.0)
3. Fatigue	3.5 (0.8)	3.8 (1.0)	3.6 (0.9)
PMSQ (1-5, 1=best)			
1. Period pain	3.4 (0.9)	3.5 (1.1)	3.5 (1.0)
2. Backache	2.8 (1.1)	3.2 (0.9)	3.0 (1.0)
3. Tender or swollen breasts	2.8 (0.9)	3.4 (1.1)	3.1 (1.0)
4. Bloating	3.5 (0.9)	3.8 (0.9)	3.7 (0.9)
5. Fluid retention	3.4 (1.0)	3.4 (1.0)	3.4 (1.0)
6. Migraine or headache	3.0 (1.2)	3.0 (1.4)	3.0 (1.3)
7. Fatigue	3.8 (0.8)	3.9 (0.9)	3.9 (0.9)
8. Irritability	3.7 (1.1)	3.6 (1.2)	3.6 (1.1)
9. Tension	3.0 (1.1)	3.5 (1.2)	3.2 (1.2)
10. Depression	2.6 (1.3)	3.5 (1.3)	3.0 (1.3)
11. Anxiety	2.4 (1.2)	3.0 (1.5)	2.7 (1.3)
12. Difficulty concentrating	2.9 (0.9)	2.9 (1.1)	2.9 (1.0)
13. Crying bouts	1.8 (0.7)	2.1 (1.2)	2.0 (1.0)
14. Clumsiness	2.6 (1.2)	2.9 (1.0)	2.7 (1.1)
15. Aggression	2.3 (1.2)	3.1 (1.4)	2.7 (1.3)

Note:
t test found no significant differences between groups

pain was 3.4, with both heavy bleeding and fatigue scoring 3.6. Women also scored the *"troublesomeness"* of peri-menstrual symptoms *over the previous three months* at baseline *[Table 29].* The mean scores for the 15 individual components of the PMSQ were below three for aggression, clumsiness, crying bouts, difficulty concentrating, and anxiety. Scores were three and above for the components period pain, backache, breast tenderness, bloating, fluid retention, headache, fatigue, irritability, tension and depression. Fatigue and pain were rated amongst the top four: Fatigue was rated the most troublesome symptom, followed by irritability (3.6), bloating (3.7) and pain (3.5). Again, t tests found no significant differences between groups.

5.2.5 Global indicators of change and menstrual health at baseline

At baseline, when asked to consider in what ways their periods had changed over the past three months, the majority (92%) answered that there had either been no change or their symptoms had worsened: 54% (21/39) reported *"no change"*, 26% (10/39) that they had *"slightly worsened"*, and 13% (5/39) that they had *"much worsened"*. A small number, 8% (3/39), reported that their periods had either *"slightly"* or *"much improved"*. And when asked how they would feel if they were told their periods would be like this until their menopause, almost all (97%, 38/39) said they would feel *"terrible"* (51%), *"unhappy"* (23%) or *"mostly dissatisfied"* (23%). Just one (3%) said that she would feel *"mixed"* – about equally satisfied or dissatisfied.

5.2.6 Avoiding medication and surgery as a treatment priority at baseline

At baseline, the majority of women reported avoiding surgery (80%, 31/39) and avoiding or cutting down medication (74%, 29/39) to be very important treatment priorities.

5.2.7 Satisfaction with 'overall care' at baseline

On treatment satisfaction and the overall care they had received in the previous three months, 35% (13/37) of participants reported being *"neither satisfied nor dissatisfied"*, 24% (9/37) reported being *"very"* or *"somewhat satisfied"*, whilst 22% (8/37) were *"somewhat"* and 19% (7/37) *"very"* dissatisfied with the care received.

Table 30: The amount of acupuncture provided in ACUMEN and co-interventions used

Practitioners	**Number of patients treated**	**Number of sessions provided per patient - range**	**Mean**	**Treatment period - months**	**Mean**	**Co-interventions**	**Advice**
A	9	19 to 20	19.8	6	6	Moxibustion, acupressure-massage, tuina, cupping, TENS, laser-acupuncture, kinesiology	Chinese dietary advice and lifestyle advice (e.g. relaxation, regular meals, appropriate exercise) and nutritional advice (e.g. iron and vitamin supplements)
B	5	15 to 20	16.8	4 to 6	5.2	Chinese herbs, acupressure-massage, moxibustion, ear acupuncture, Bowen technique	Chinese dietary advice and lifestyle advice (e.g. relaxation, regular meals, appropriate exercise) and nutritional advice (e.g. acidophilus and Echinacea supplements)
C	6	10 to 20	16.7	4 to 6	5.2	None	Chinese dietary advice and lifestyle advice (e.g. relaxation, regular meals, appropriate exercise)
Total	20	343	17.1	111	5.5		

5.3 *The interventions: acupuncture in addition to usual care – 'experimental' versus usual care alone – 'comparator'*

5.3.1 *Compliance with acupuncture treatment*

All 20 of the women randomised to and offered the option of attending for acupuncture accepted this offer, and all 20 completed their course of up to 20 sessions of acupuncture within a six month treatment period *[Table 30]*. The number of treatments received ranged from 10 to 20, with 14 (70%) of patients receiving 18 to 20 sessions. The range of co-interventions used by the practitioners in ACUMEN is given in Table 30.

5.3.2 *Treatment over the course of ACUMEN*

The use of medical contraceptives, drug therapies, nutritional supplements and CAM modalities at baseline and over the course of ACUMEN is summarised in Table 27-drugs. Chi-squared tests found no significant difference between groups in treatments used, although both women using the intra-uterine contraceptive device (IUCD) were in the usual care group.

There were few tests and referral for investigations after joining ACUMEN. At three months two women in the acupuncture group and one in the usual care group had received blood tests to test for anaemia, and one woman in the acupuncture group had undergone pelvic ultrasound to investigate the cause of her menorrhagia. At six months one woman in the usual care group had undergone hysteroscopy and one had been put on the waiting list for hysterectomy. And at 12 months one woman in the acupuncture group had been put on the waiting list for endometrial ablation.

5.3.3 *Side effects, reactions and adverse events at three, six and 12-months follow-up*

At 3 months the possible side-effects from drug therapies or reactions to acupuncture experienced by participants were very similar in both groups; the exception was *"feeling drowsy"*, reported by 58% (10/17) of respondents in the acupuncture group as opposed to 17% (2/12) in the usual care group (chi squared = 5.2; $p < 0.05$). At six months, *"feeling drowsy"* was balanced between groups, and instead *"stomach pain"* appeared more of a concern for those in the usual care group (75%) than the acupuncture group (28%, 5/18) [chi squared = 6.5;

p ~ 0.01]. Whilst at 12-month follow-up *"breast tenderness"* (73%, 8/11) and *"depression"(46%, 5/11)* appeared more of a problem for the usual care group than the acupuncture group (50%, 9/18; 22%, 4/18 respectively), with skin problems more of a problem for those in the acupuncture group (22%, 4/18) than in the usual care group (9%, 1/11).

Eighteen of the 20 women receiving acupuncture as part of ACUMEN reported on three types of adverse events *[Table 31]*: (1) pain on needle insertion, (2) bruising from needle insertion, and (3) bleeding for more than ten seconds because of needle insertion. Pain on insertion of the needle is very common and bruising quite common, as is bleeding for a few seconds – but bleeding for more than ten seconds is not. In the ACUMEN patient treatment booklets, the acupuncture practitioners recorded a total of eight adverse events: four incidents of bruising, one 'forgotten needle', two incidents of skin reactions to the needle, and one incident of bleeding for more than ten seconds: The 'forgotten needle' at acupoint DU20 was safely removed by the patient before leaving the treatment room and informing the practitioner; one skin reaction involved a *"slight swelling and reddening at (acupoint) P5"* that had lessened before the patient's leaving the treatment room; the other was described as *"tiny marks at (acupoints) Liv3 and SP6 as a result of a nickel allergy"* – a response that was reported at the patients seventeenth treatment only. None of the patients reported the skin reactions or the forgotten needle.

Table 31: Adverse events reported during the ACUMEN acupuncture treatment period

Adverse event	**Patient reports at 3 months** (max n=20/20)	**Patient reports at 6 months** (max n=20/20)	**Practitioner reports in the ACUMEN treatment booklets – total number**
1. Pain where acupuncture needle inserted	67% (12/18)	56% (10/18)	0
2. Bruising where acupuncture needle inserted	24% (4/17)	29% (5/17)	4
3. Bleeding > 10 seconds where acupuncture needle inserted	6% (1/18)	24% (4/17)	1
4. Skin reaction at site	0	0	2
5. Forgotten needle	0	0	1

5.4 Differences in primary outcomes attributable to acupuncture

5.4.1 The AMQ – Differences in scores at three, six and 12 months

At three months, analysis of covariance revealed how both groups had improved *[Table 32]*. The acupuncture (in addition to usual care) group had improved by more (6.9), falling from a "severe" GP-assessed menorrhagia severity score at baseline to a "very mild" score, while the usual care group fell to a "mild" score. But the difference between groups was not significant and the confidence interval crossed zero (-5.0 to 18.8). The null hypothesis that the error of variance of the dependent variable is equal across groups was not violated (p=0.818). By six-months the acupuncture group had improved significantly more than the usual care group (p=0.015); whilst both groups achieved a "mild" GP-assessed menorrhagia severity score, there was a difference of 12.6 points between groups, equivalent to a 'very large' effect size of more than 1.0. By 12 months, both groups had deteriorated, but had remained around the level for a "mild" GP-assessed menorrhagia severity score. The acupuncture group continued to be better off than the usual care group, but the difference between groups (6.2 points) was no longer significant, which may reflect reduced power from loss to follow-up (n=32/39).

5.4.2 The SF-36 – Differences in scores at three, six and 12 months

PCS and MCS

At three months, there was little change in the physical component score (PCS) and mental component score (MCS) for the usual care group, and the improvements in the acupuncture group were not significant *[Table 32]*. Both groups had improved again by six months and those in the acupuncture group continued to be better off, but again the differences were not significant. By 12 months follow-up PCS scores for the acupuncture group had deteriorated and become very similar between groups, whilst the difference between groups on the MCS had widened with a difference of nearly ten points to achieve statistical significance at the 0.05 level.

Eight dimensions

At three months, the acupuncture group had improved on all eight individual dimensions and improvements were significant at the 0.05 level for social functioning, vitality and pain *[Table 32]*. However, by six months only differences in the dimension pain remained

Table 32: Differences in primary outcomes attributable to acupuncture and corrected for baseline values of outcome measure in question[a] - [values are mean (SE) scores unless stated otherwise]

	At three months			At six months			At 12 months		
Outcome measure	**Acupuncture + usual care** (max n=20/20)	**Usual care** (max n=17/19)	**Net benefit of acupuncture[b] (95% CI)**	**Acupuncture + usual care** (max n=20/20)	**Usual care** (max n=16/19)	**Net benefit of acupuncture[b] (95% CI)**	**Acupuncture + usual care** (max n=18/20)	**Usual care** (max n=14/19)	**Net benefit of acupuncture[b] (95% CI)**
AMQ (0-100, 0=best)	33.5 (3.9)	40.4 (4.2)	+ 6.9 (-5.0 to +18.8)	26.0 (3.2)	38.5 (3.6)	**+12.6** (+2.6 to + 22.5)**	31.5 (4.1)	37.7 (4.7)	+6.2 (-6.9 to +19.2)
SF-36 (mean=50, SD=10, 100=best)									
PCS – Physical Component Score	46.4 (1.3)	43.6 (1.4)	+2.8 (-1.0 to +6.6)	50.0 (1.5)	46.8 (1.6)	+3.2 (-1.3 to +7.7)	48.2 (1.8)	47.7 (2.0)	+0.5 (-5.0 to +5.9)
MCS – Mental Component Score	48.2 (2.6)	41.6 (2.9)	+6.6 (-1.3 to +14.7)	46.8 (3.0)	42.7 (3.2)	+4.0 (-5.0 to +13.1)	49.0 (2.7)	39.54 (3.01)	**+9.4* (+1.1 to +17.8)**
1 Physical functioning	81.7 (3.9)	79.4 (4.2)	+2.3 (-9.5 to +14.1)	82.8 (4.1)	83.6 (4.48)	-0.8 (-13.2 to +11.6)	84.8 (4.0)	83.1 (4.5)	+1.6 (-10.7 to +14.0)
2 Social functioning	77.9 (5.1)	61.3 (5.5)	**+16.6* (+1.2 to +31.9)**	78.5 (5.0)	69.9 (5.6)	+8.6 (-6.9 to +24.0)	75.2 (6.6)	64.1 (7.4)	+11.1 (-9.4 to +31.5)
3 Role physical	73.0 (5.3)	62.3 (5.7)	+10.7 (-5.2 to +26.6)	77.1 (4.3)	71.4 (4.7)	+5.7 (-7.3 to +18.7)	78.6 (4.6)	69.9 (5.2)	+8.8 (-5.5 to +23.0)
4 Role mental	82.8 (6.2)	71.6 (6.7)	+11.2 (-7.4 to +29.8)	78.8 (5.9)	76.2 (6.4)	+2.5 (-15.2 to +20.2)	83.0 (5.2)	68.9 (5.8)	+14.1 (-1.9 to +30.1)
5 Mental health	73.7 (4.2)	62.0 (4.6)	+11.7 (-1.0 to +24.4)	71.8 (4.6)	63.2 (5.0)	+8.5 (-5.4 to +22.5)	75.0 (4.3)	63.9 (4.9)	+11.2 (-2.2 to +24.6)
6 Vitality	54.6 (4.5)	39.0 (4.9)	**+15.6* (+2.0 to +29.2)**	56.6 (5.2)	43.7 (5.7)	+13.0 (-2.8 to +28.8)	57.0 (4.6)	37.5 (5.3)	**+19.4** (+5.0 to +33.8)**
7 Pain	62.2 (4.3)	38.3 (4.6)	**+23.9** (+11.0 to +36.8)**	76.2 (3.6)	46.2 (4.0)	**+30.0** (+19.0 to +41.1)**	61.2 (5.6)	54.1 (6.4)	+7.2 (-10.4 to +24.7)
8 General health	64.2 (4.4)	55.0 (4.8)	+9.3 (-4.0 to +22.5)	67.6 (4.1)	59.5 (4.6)	+8.1 (-4.5 to +20.8)	66.4 (4.8)	56.0 (5.3)	+10.4 (-4.4 to +25.1)
			9/9			*8/9*			*9/9*

**Mean difference is significant at the .05 level*
*** Mean difference is significant at the .01 level*

Table 32: Differences in primary outcomes attributable to acupuncture & corrected for baseline values of outcome measure in question

Notes to accompany the table

a It is traditional to analyse longitudinal studies like ACUMEN by calculating changes between baseline and follow-up scores. However Vickers & Altman (BMJ 2001;323;1123-4) have shown that this is statistically less efficient than using analysis of covariance to correct follow-up scores for the corresponding baseline scores of the outcome measure in question. This chapter therefore uses analysis of covariance rather than changes in scores.

b While the AMQ has a best score of 0, the SF-36 has a best score of 100. To ensure that this table presents the 'net benefits of acupuncture' consistently, the 4th, 7th and 10th columns present genuine gains from acupuncture with positive signs and losses with negative signs.

c All nine of the ostensibly independent 'net benefits of acupuncture' (i.e. excluding PCS and MCS, which together summarise the eight dimensions of Sf-36) after three months are positive; eight out of nine after six months are positive; and nine out of nine after 12 months are positive. According to the Sign Test (Armitage & Berry. *Statistical methods in medical research*, 2nd ed. Oxford: Blackwell; 1987: 409) these three findings are all significant – at the 1%, 5% and 5% level respectively.

d Despite the sample size of only 39, six of the 27 t tests (again excluding PCS and MCS) in this table summarised by confidence intervals are significant at the 0.05 level (marked by asterisks and bold type) – more than the 1.4 expected at random. Moreover, of these six, four are significant at the 0.01 level (marked by two asterisks and bold type) – even 1 spurious significance at this level is unlikely.

statistically significant, with a difference of more than 30 points between groups. And whilst a difference of 13 points was observed for the dimension vitality, this did not achieve statistical significance again until 12 months.

5.4.3 Benefits from acupuncture

In summary, despite the maximum sample size of only 39, six of the 27 t tests in Table 32 represented by confidence intervals are significant at the 0.05 level. This is much more than the 1.4 that could be expected at random (viz 5% of 27). Moreover, of these six, four are significant at the 0.01 level, which is most notable, as even 1 spurious significance is unlikely at this level. Furthermore, the effect sizes for pain and the AMQ at six months are especially large – 1.22 and 1.03 respectively.

The findings in Table 32 can also be summarised by focusing on the direction of change. All nine of the ostensibly independent 'net benefits of acupuncture' (i.e. excluding PCS and MCS, which together summarise the eight dimensions of the SF-36, and excluding the EQ-5D, which is also a summary health measure) after three months were positive; eight out of nine after six months were positive; and nine out of nine after 12 months were positive. According to the Sign Test (Armitage & Berry 1987) these three combinations are all significant – at the 1%, 5% and 5% level respectively. The implications of all these promising findings need further discussion.

5.4.4 Area under the curve analysis for the AMQ and SF-36

Thus analysis of AMQ and SF-36 scores leads to six significant t tests out of 27, and three significant Sign Tests out of three. Because the sample size was at most 39, random variation was capable of swamping even moderate to large changes in the AMQ and SF-36 scores. So an 'area under curve analysis' was undertaken to estimate the total benefit from acupuncture over 12 months. Taking the average scores at baseline, three, six and 12-months, and giving proportionately more weight to the six-month and 12-month scores (because six months had elapsed between these measures) reduces the effect of individual scores. Nevertheless, this analysis did not find the differences between groups statistically significant.

5.5 Differences in secondary outcomes attributable to acupuncture

5.5.1 YMRP and PMSQ – Differences in scores at three, six and 12-months follow-up

York Menorrhagia-Relevant Profile - YMRP

By three months, both groups had improved scores on all three dimensions – heavy bleeding, period pain and fatigue *[Table 33]*. Analysis of covariance found the acupuncture group significantly better off at the level of 0.05 across all dimensions at six months. Whilst the acupuncture group continued to fare better at 12 months, only differences in bleeding continued to show statistical significance. However, all three dimensions showed gains from acupuncture at each follow-up, as depicted by the positive signs in the fourth, seventh and tenth columns of Table 33.

Peri-menstrual Symptoms Questionnaire – PMSQ

At three, six and 12 months, analysis of covariance found gains from acupuncture across all 15 PMSQ dimensions, as depicted by the positive signs in the fourth, seventh and tenth columns of Table 34. The mean difference between groups was significant at the 0.05 level for six of the dimensions at three-months, ten of the dimensions at six-months, and nine of the dimensions at 12-months follow-up. All 45 net differences favoured acupuncture – a highly significant finding on the Sign Test.

5.5.2 Global indicators of change and menstrual health at three, six and 12.months

At three months 90% (18/20) of the acupuncture group reported *"slight"*, *"moderate"* or *"vast improvement"* in their periods compared with only 35% (6/17) of the usual care group [chi squared = 12.1; $p < 0.001$]. And 50% (10/20) of the acupuncture group reported that they would be *"mostly satisfied"* (15%, 3/20), *"pleased"* (20%, 4/20) or *"delighted"* (15%, 3/20) if their periods remained like this until their menopause. Conversely, in the usual care group, only three (18%) reported that they would be *"mostly satisfied"*, *"happy"* or *"delighted"* [chi squared = 4.2; $p < 0.05$].

At six months, the end of the acupuncture treatment period, 44% (7/16) of the usual care group and 95% (19/20) of the acupuncture group reported *"improvement"* in their periods [chi squared = 7.7; $p < 0.01$]. One patient in the acupuncture group (5%) and two in the usual

Table 33: Differences in key symptoms at patients' most recent period (YMRP) attributable to acupuncture and corrected for baseline values – [values are mean (SE) scores]

	three months			six months			12 months		
Outcome measure	**Acupuncture + Usual care** (max n=20/20)	**Usual care** (max n=17/19)	**Net benefit of acupuncture (95% CI)**	**Acupuncture + Usual care** (max n=20/20)	**Usual care** (max n=16/19)	**Net benefit of acupuncture (95% CI)**	**Acupuncture + Usual care** (max n=18/20)	**Usual care** (max n=14/19)	**Net benefit of acupuncture (95% CI)**
YMRP (1-5, 1=best)									
1. Heavy bleeding	2.6 (0.3)	3.2 (0.3)	+0.6 (-0.3 to 1.5)	2.2 (0.3)	3.3 (0.3)	**+1.2* (0.2 to 2.1)**	2.1 (0.3)	3.0 (0.3)	**+0.9* (0.1 to 1.7)**
2. Period pain	2.0 (0.3)	3.3 (0.3)	**+1.3* (0.5 to 2.1)**	2.0 (0.3)	3.3 (0.3)	**+1.3* (0.6 to 2.1)**	2.4 (0.3)	3.0 (0.3)	+0.6 (-0.4 to 1.6)
3. Fatigue	2.7 (0.3)	3.3 (0.3)	+0.6 (-0.3 to 1.4)	2.4 (0.3)	3.4 (0.3)	**+1.0* (0.2 to 1.7)**	2.4 (0.2)	3.1 (0.3)	+0.6 (-0.2 to 1.4)

**Mean difference is significant at the .05 level*

Table 34: Differences in peri-menstrual symptoms attributable to acupuncture and corrected for baseline values – [values are mean (SE) scores unless stated otherwise]

	At three months			At six months			At 12 months		
Outcome measure	**Acupuncture + Usual care** (max n=20/20)	**Usual care** (max n=17/19)	**Net benefit of acupuncture (95% CI)**	**Acupuncture + Usual care** (max n=20/20)	**Usual care** (max n=16/19)	**Net benefit of acupuncture (95% CI)**	**Acupuncture + Usual care** (max n=18/20)	**Usual care** (max n=14/19)	**Net benefit of acupuncture (95% CI)**
PMSQ (1-5, 1=best)									
1. Period pain	2.4 (0.2)	3.3 (0.2)	**+0.9* (0.2 to 1.6)**	2.1 (0.2)	3.0 (0.2)	**+0.9* (0.3 to 1.6)**	2.6 (0.3)	3.0 (0.3)	+0.4 (-0.4 to 1.2)
2. Backache	2.29 (0.2)	2.9 (0.2)	**+0.6* (0.005 to 1.3)**	2.0 (0.2)	2.9 (0.2)	**+0.9* (0.4 to 1.5)**	2.1 (0.2)	2.7 (0.2)	+0.5 (-0.05 to 1.1)
3. Breast tenderness	1.9 (0.2)	2.5 (0.2)	**+0.6* (0.007 to 1.2)**	1.8 (0.2)	2.7 (0.2)	**+0.9* (0.3 to 1.4)**	2.0 (0.2)	2.6 (0.2)	**+0.7* (0.05 to 1.4)**
4. Bloating	2.9 (0.2)	3.0 (0.3)	+0.7 (-0.01 to 1.4)	1.8 (0.2)	3.1 (0.2)	**+1.2* (0.7 to 1.8)**	2.1 (0.2)	3.1 (0.2)	**+1.0* (0.3 to 1.6)**
5. Fluid retention	2.1 (0.3)	2.7 (0.3)	+0.6 (-0.1 to 1.4)	1.8 (0.2)	3.0 (0.2)	**+1.2* (0.6 to 1.8)**	1.9 (0.2)	2.8 (0.2)	**0.9* (0.2 to 1.6)**
6. Migraine or headache	1.7 (0.2)	2.7 (0.2)	**+1.0* (0.4 to 1.6)**	1.8 (0.2)	2.7 (0.3)	**+0.9* (0.2 to 1.6)**	2.0 (0.2)	2.3 (0.3)	+0.3 (-0.4 to 1.1)
7. Fatigue	2.8 (0.3)	3.4 (0.3)	+0.6 (-0.2 to 1.4)	2.4 (0.2)	3.4 (0.2)	**+1.0* (0.4 to 1.7)**	2.7 (0.2)	3.4 (0.2)	**+0.7* (0.1 to 1.3)**
8. Irritability	2.1 (0.2)	3.2 (0.3)	**+1.1* (0.4 to 1.8)**	1.8 (0.2)	3.4 (0.3)	**+1.6* (0.9 to 2.3)**	2.3 (0.2)	3.4 (0.3)	**+1.0* (0.3 to 1.8)**
9. Tension	2.0 (0.3)	2.6 (0.3)	+0.6 (-0.3 to 1.2)	1.9 (0.3)	3.2 (0.3)	**+1.3* (0.5 to 2.1)**	2.2 (0.3)	3.0 (0.3)	**+0.9* (0.04 to 1.7)**
10. Depression	2.0 (0.3)	2.3 (0.3)	+0.4 (-0.3 to 1.2)	1.8 (0.2)	2.8 (0.3)	**+1.0* (0.3 to 1.7)**	1.9 (0.3)	2.9 (0.3)	**+1.0* (0.1 to 1.8)**
11. Anxiety	1.7 (0.3)	2.0 (0.3)	+0.3 (-0.6 to 1.2)	1.7 (0.3)	2.4 (0.3)	+0.7 (-0.1 to 1.4)	1.6 (0.3)	2.2 (0.3)	+0.6 (-0.2 to 1.5)
12. Difficulty concentrating	2.2 (0.3)	2.3 (0.3)	+0.05 (-0.7 to 0.8)	2.0 (0.3)	2.5 (0.3)	+0.4 (-0.5 to 1.3)	2.0 (0.3)	2.6 (0.3)	+0.7 (-0.2 to 1.5)
13. Crying bouts	1.2 (0.2)	2.0 (0.2)	**+0.8* (-0.3 to 1.3)**	1.3 (0.2)	1.7 (0.2)	+0.4 (-0.1 to 0.8)	1.3 (0.2)	1.9 (0.2)	**+0.6* (0.02 to 1.2)**
14. Clumsiness	1.9 (0.2)	2.2 (0.3)	+0.3 (-0.4 to 1.0)	1.8 (0.2)	2.1 (0.3)	+0.2 (-0.5 to 1.0)	1.8 (0.2)	2.2 (0.3)	+0.4 (-0.3 to 1.2)
15. Aggression	1.7 (0.2)	2.3 (0.3)	+0.6 (-0.1 to 1.4)	1.7 (0.3)	2.3 (0.3)	+0.6 (-0.2 to 1.4)	1.6 (0.2)	2.3 (0.2)	**+0.8* (0.2 to 1.3)**

*Mean difference is significant at the .05 level

care group (12%) reported worsening of symptoms. Asked how they would feel if their periods remained like this until their menopause, the majority of the acupuncture group (75%, 15/20) said they would feel *"mostly satisfied"* (15%, 3/20), *"pleased"* (20%, 4/20) or *"delighted"* (40%, 8/20), as compared with 19% (3/16) in the usual care group [chi squared = 11.2; $p < 0.001$].

At 12 months, 11 (61%) patients in the acupuncture group reported that, in the six months since the end of the acupuncture treatment period, their periods had improved; three (17%) women's periods had remained unchanged; and four (22%) reported worsening of symptoms. In comparison, 36% of women in the usual care group (5/14) reported an improvement in their periods, 36% reported *"no change"* and four (28%) reported that they had *"worsened"*. Given the significantly better initial improvement in the acupuncture group, it is not surprising that their subsequent improvement was not significant [chi squared = 2.0]. If told their periods would remain unchanged until their menopause, 44% (8/18) of the acupuncture group said they would feel *"mostly satisfied"* (6%, 1/18), *"pleased"* (22%, 4/18) or *"delighted"* (17%, 3/18), as compared with 21% (3/14) in the usual care group. Four women (22%) in the acupuncture group and two (14%) in the usual care group said they would feel about equally satisfied and dissatisfied. And most of the women in the usual care group (64%) and a third (33%) of the women in the acupuncture group said that they would feel *"dissatisfied"*, *"unhappy"* or *"terrible"* about this [chi squared = 1.8; not statistically significant].

5.6 Patients' treatment priorities, satisfaction and preferences

5.6.1 Avoiding medication and surgery as priorities at three, six and 12-months follow-up

At three months, all participants considered cutting down or avoiding medication to be a treatment priority. However, only 59% (10/17) of the usual care group in comparison with 90% (18/20) of those in the acupuncture group, considered this a *"very important"* treatment priority [chi squared = 4.9; $p < 0.05$]. Again 59% of the usual care group and 90% in the acupuncture group considered avoiding surgery *"very important"* [chi squared = 4.9; $p < 0.05$]; however one woman in each group reported that she did not consider this an important treatment priority.

At six months most, but not all participants (92%, 33/36), considered avoiding or cutting down medication to be an important treatment priority; the proportion considering this *"very important"* was 65% (13/20) in the acupuncture group and 50% (8/16) in the usual care group [chi squared = 0.8; not statistically significant]. Three quarters (12/16) of the usual care group and 90% (18/20) of the acupuncture group considered avoiding surgery *"very important"* [chi squared = 1.4; not statistically significant]; again, one woman in each group did not consider this a treatment priority. By 12 months, 19 of 20 women responding to these two questions reported that reducing medication and avoiding surgery were treatment priorities.

5.6.2 Satisfaction with care at three, six and 12-months

At three months, 89% (17/19) of those in the acupuncture group and 62% (8/13) of those in the usual care group reported being *"very"* or *"somewhat satisfied"* with the overall care they had received in the previous three months [chi squared = 3.5; not statistically significant], with one patient in the usual care group reporting that they had been *"somewhat dissatisfied"*. At six months, 95% (18/19) of those in the acupuncture group reported being *"very satisfied"* with the care they had received since joining the trial, compared with only 55% (8/11) of those in the usual care group [chi squared = 7.0; $p < 0.01$]. Again, just one patient in the usual care group reported that they had been *"dissatisfied"*. At 12 months, 75% of both the eight responders in the acupuncture group, and the four responders in the usual care group, stated that they had been *"very"* or *"somewhat satisfied"* [chi squared = 0.0; not significant]; none of the twelve reported being dissatisfied with care in the previous six months.

5.6.3 Treatment preferences and acceptability

At six-months, the end of the acupuncture treatment period, all 20 women randomised to and accepting the offer of acupuncture *"strongly agreed"* or *"agreed"* that if they went back six months they would make that same decision to use acupuncture *[Table 35]*. Those in the usual care group were less certain they would use prescribed medication, with 63% (10/16) reporting they would want instead to have acupuncture [chi squared = 9.0; $p < 0.001$]. And the majority of women in both groups (85%, 28/33) agreed they would not choose a

Table 35: Treatment preferences and acceptability [with most popular response category emboldened]

Knowing what I know now about treatments for heavy periods.....		Six month follow-up		12-month follow-up	
		Acupuncture group (max n=20/20)	Usual care group (Max n=16/19)	Acupuncture group (max n=20/20)	Usual care group (Max n=16/19)
If I could TRAVEL BACK in time I would have ACUPUNCTURE	Strongly disagree			6% (1/17)	
	Disagree				
	Uncertain		38% (6/16)		**46% (6/13)**
	Agree	20% (4/20)	13% (2/16)	18% (3/17)	15% (2/13)
	Strongly agree	**80% (16/20)**	**50% (8/16)**	**77% (13/17)***	39% (5/13)*
If I could TRAVEL BACK in time I'd have MEDICATION	Strongly disagree	**59% (10/17)**	**31% (5/16)**	33% (5/15)	**39% (5/13)**
	Disagree	18% (3/17)	25% (4/16)	33% (5/15)	15% (2/13)
	Uncertain	12% (2/17)	19% (3/16)	13% (2/15)	31% (4/13)
	Agree	12% (2/17)	19% (3/16)	13% (2/15)	8% (1/13)
	Strongly agree		6% (1/16)	7% (1/15)	8% (1/13)
If I could TRAVEL BACK in time I'd have a HYSTERECTOMY	Strongly disagree	**77% (13/17)**	**81% (13/16)**	**83% (10/16)**	**69% (9/13)**
	Disagree	6% (1/17)	6% (1/16)	19% (3/16)	15% (2/13)
	Uncertain	12% (2/17)	6% (1/16)		8% (1/13)
	Agree	6% (1/17)	6% (1/16)	13% (2/16)	8% (1/13)
	Strongly agree			6% (1/16)	
If in the FUTURE I needed treatment I'd want ACUPUNCTURE	Strongly disagree				
	Disagree			6% (1/17)	
	Uncertain	5% (1/20)	19% (3/16)	6% (1/17)	**46% (6/13)**
	Agree	15% (3/20)	25% (4/16)	12% (2/17)	15% (2/13)
	Strongly agree	**80% (16/20)**	**56% (9/16)**	**77% (13/17)***	39% (5/13)*

Table 35 continued: Treatment preferences and acceptability [with most popular response category emboldened]

If in the FUTURE I needed treatment I'd want MEDICATION	Strongly disagree	**53% (9/17)**	**38% (6/16)**	**47% (7/15)**	**31% (4/13)**
	Disagree	24 (4/17)	19% (3/16)	20% (3/15)	23% (3/13)
	Uncertain	12% (2/17)	25% (4/16)	20% (3/15)	31% (4/13)
	Agree	12% (2/17)	13% (2/16)	7% (1/15)	8% (1/13)
	Strongly agree		6% (1/16)	7% (1/15)	8% (1/13)
If in the FUTURE I needed treatment I'd want a HYSTERECTOMY	Strongly disagree	**81% (13/16)**	**88% (14/16)**	**50% (8/16)**	**70% (9/13)**
	Disagree	6% (1/16)		19% (3/16)	15% (2/13)
	Uncertain	13% (2/16)	6% (1/16)	13% (2/16)	8% (1/13)
	Agree		6% (1/16)	13% (2/16)	8% (1/13)
	Strongly agree			6% (1/16)	
I'd recommend ACUPUNCTURE to other women	Strongly disagree			11% (2/18)	
	Disagree			6% (1/18)	7% (1/14)
	Uncertain	5% (1/20)	53% (8/15)	6% (1/18)	**57% (8/14)**
	Agree	10% (2/20)	13% (2/15)	6% (1/18)	21% (3/14)
	Strongly agree	**85% (17/20)****	**33% (5/15)****	**72% (13/18)****	14% (2/14)**
I'd recommend MEDICATION to other women	Strongly disagree	**35% (6/17)**	**33% (5/15)**	**50% (7/14)**	**36% (5/14)**
	Disagree	24% (4/17)	20% (3/15)	21% (3/14)	7% (1/14)
	Uncertain	29% (5/15)	27% (4/15)	14% (2/14)	43% (6/14)
	Agree	12% (2/15)	13% (2/15)	14% (2/14)	7% (1/14)
	Strongly agree		7% (1/15)		7% (1/14)
I'd recommend a HYSTERECTOMY to other women	Strongly disagree	**77% (13/17)**	**60% (9/15)**	**54% (7/13)**	**50% (7/14)**
	Disagree	12% (2/17)	27% (4/15)	23% (3/13)	7% (1/14)
	Uncertain	12% (2/17)	6% (1/16)	15% (2/13)	43% (6/14)
	Agree			8% (1/13)	
	Strongly agree				

After combining adjacent categories 2 of 18 2x2 chi-squared tests show difference between acupuncture & usual care significant at 1% level and another 2 show difference significant at the 5% level

hysterectomy. If they were to need further treatment for menorrhagia in the future, 95% (19/20) of the acupuncture group and 81% (13/16) of the usual care group would choose acupuncture. Again the great majority (84%, 27/32) would not choose hysterectomy. When asked whether they would recommend acupuncture to other women, 95% (19/20) of those in the acupuncture group said that they would, whilst the majority in both groups stated that they would not recommend prescribed medications (56%, 18/32) or hysterectomy (87%, 28/32) to others. At 12-months responses to these items were similar to those at six months: women continued to show preference for acupuncture over medication and surgery, with those in the acupuncture group exhibiting stronger preferences.

6. Discussion

6.1 Justification for a full-scale definitive trial

Despite the maximum sample size of only 39, the findings from the ACUMEN trial are most promising. At three months, participants randomised to the offer of acupuncture in addition to usual care from their GP reported a greater improvement on the Aberdeen Menorrhagia Questionnaire than the usual care group, falling from a "severe" GP-assessed menorrhagia severity score at baseline to a "very mild" score, while the usual care group fell to a "mild" score. By six-months the difference between groups was 12 points, equivalent to a very large effect size of more than 1.0. The acupuncture group continued to be better off then the usual care group at 12-months, but the difference of six points between groups was no longer significant. This may reflect reduced power from loss to follow-up (n=32/39).

Evidence of genuine gains from acupuncture also came in the differences in scores on the SF-36. At three and six months the acupuncture group had improved by more than usual care group on the physical (PCS) and mental component scores (MCS), and at 12-month follow-up there was a statistically significant difference of nearly 10 points on the MCS. The acupuncture group had improved by more than the usual care group on all eight individual dimensions at three months, and improvements were significant for the dimensions social functioning, vitality and pain. By six months only differences in the dimension pain remained significant, with an especially large difference of 30 points between groups. And whilst a difference of 13 points was observed for the dimension vitality, this did not achieve statistical significance again until 12-months.

Thus six of the 27 t tests of primary outcomes are significant at the 0.05 level – far more than the 1.4 that could be expected at random (viz 5% of 27). Moreover, of these six, four are significant at the 0.001 level, and even one spurious significance at this level is unlikely. Focusing on the direction of differences, all nine of the ostensibly independent 'net benefits of acupuncture' (i.e. excluding PCS and MCS which are summary measures) after three months were positive; eight out of nine after six-months were positive; and nine out of nine after 12 months were positive. All three of these combinations are significant at the 1%, 5% and 1% level respectively.

These findings were reinforced by the secondary outcomes of ACUMEN. First, three outcomes devised for ACUMEN focus upon symptoms most relevant to patients with menorrhagia – heavy bleeding, pain and fatigue – thus estimating the wider effects of treatment upon menstrual health. The acupuncture group consistently reported less troublesome symptoms, with five of nine statistical tests at three, six and 12 months significant at the level of 0.05. Second, women in the acupuncture group reported fewer and less severe peri-menstrual symptoms, particularly period pain, backache, breast tenderness, bloating, fluid retention, irritability, headache, fatigue, depression, tension, aggression and crying bouts, all of which achieved statistical significance at least once over the course of ACUMEN; in all 25 statistical tests out of 45 were significant. Taken together, all 54 net gains from acupuncture across all 18 condition-specific dimensions at all three times were positive – a very convincing finding! Third, on the global indicator that asked how their periods had changed, at three months 90% and at six months 95% of women in the acupuncture group reported improvements, compared with only 35% in the usual care group at three months and 44% at six months – both significant differences. When asked how they would feel if their periods remained like this until the menopause, 50% of the acupuncture group at three months and 75% at six months reported satisfaction with treatment outcome, compared with 18% of the usual care group at three months and 19% at six months – two more significant differences. However, satisfaction with treatment outcome dropped back to 44% within the acupuncture group at 12 months. This reflects the deterioration in menstrual health on the Aberdeen Menorrhagia Questionnaire and the pain dimensions of both the SF-36 and York Menorrhagia-Relevant Profile.

Taken together, these findings provide convincing evidence of genuine effects and justify a full-scale definitive trial, especially when taken alongside the evidence of patients' preferences for acupuncture and their satisfaction with treatment over the course of ACUMEN. For it is notable that the great majority (92%) of patients initiated their own referral, as they wished to use acupuncture. Most women reported at baseline that reducing or avoiding medication (74%) and surgery (80%) were *"very important"* treatment priorities; a finding that is consistent with other studies of CAM users (Barrett et al 2003; Luff & Thomas 2000; Gamon 2000; Paterson & Britten 1999; Cassidy 1998a). ACUMEN also provides evidence of treatment acceptability with 100% compliance rates, and a significant difference in satisfaction between groups by the end of the acupuncture treatment period: 95% of those in the acupuncture group compared with 55% of those in the usual care group reported being *"very satisfied"* with the care they had received. Moreover, at the end of the acupuncture treatment period, all twenty of the women who had received acupuncture in ACUMEN agreed that if they travelled back in time they would make that same decision to use acupuncture. Nearly two thirds of the usual care group reported that they would choose acupuncture in retrospect. When asked whether they would recommend acupuncture to other women, 95% of those in the acupuncture group said that they would, whilst the majority of both groups would not recommend prescribed medications (56%) or hysterectomy (87%). The response patterns at 12-months were similar. These findings support the evidence of benefit from acupuncture and reduce the possibility of erroneous positive findings – Type 1 error.

6.2 Feasibility of a full-scale definitive trial

6.2.1 Participant recruitment and implications for a full-scale trial

ACUMEN achieved 39 of a target of 40 participants by responding swiftly to problems with recruitment; initially largely dependent upon the willingness of GPs to refer suitable patients, ACUMEN introduced two further recruitment pathways that were patient centred, and asked more GPs for support beyond the initial target of 10. By the end of the extended 11-month recruitment period a total of 43 GPs had agreed to facilitate referral for willing patients; 35 patients participated in ACUMEN in response to patient centred strategies, and only four on the invitation of their GP during a consultation. Even so the recruitment rate was just 3.6 patients per month. These findings have serious implications for a full trial reliant upon GPs, as it is not unusual for trials to be compromised or even fail on account of poor recruitment

rates. The postal survey of GPs in the York and Selby Primary Care Trust reported in Chapter Four was designed to understand the barriers to active GP participation in ACUMEN and assess how they might be overcome.

6.2.2 Implications of recruitment strategies upon effect size and the trial design

All but four of the 39 women participating in ACUMEN came via patient centred recruitment strategies, and all but one would have chosen to receive acupuncture if given the opportunity. This raises questions about the effect of preferences upon outcomes. Whilst the placebo effect of patient preference was maximised for those randomised to their preferred option of acupuncture within normal NHS management, for those randomised to their less preferred option of usual care from their GP alone, the placebo effect was in essence minimised, or at least greatly reduced. This has implications for the design of any full-scale trial.

An alternative pragmatic design described by Thomas and Fitter (2002) to address the problem of patients being unwilling to be randomised to usual care uses partial randomisation, where patients who express a preference are allocated to the treatment that they prefer and patients who express no preference are randomised between treatment arms. However, as Thomas and Fitter acknowledge, whilst this design estimates the effect of receiving the treatment of choice, it does not answer the more practical question of how a population of patients would fare in conditions where treatment was offered within normal NHS management. Moreover, if patient centred strategies facilitate the primary recruitment pathway in a full trial, just as they did in ACUMEN, the trial would fail to recruit sufficient numbers to three of the four arms of the trial: randomised to acupuncture; randomised to usual care; and preference for usual care. Indeed, Grimm (2002 p.149) states that, "*in large trials, advertising and mailings usually yield the largest number of randomised participants.*"

The way in which patients came to be involved in the study also has implications for the generalisability of the findings. It is of note that almost half (19/39) had previously consulted a CAM practitioner and that 36% (14/39) had previously used acupuncture, with six of the women currently receiving acupuncture and recruited by their acupuncturist. Such findings suggest that the women in ACUMEN may differ in important ways to a usual GP population of patients with menorrhagia. Analysis of demographic characteristics compared with 2001

Census data might answer questions about their representativeness and is an area of future research planned by the candidate.

6.2.3 Response rates and loss to follow-up

ACUMEN benefited from response rates of 95% (37/39), 92% (36/39) and 82% (32/39) at three, six and 12 months respectively. These rates exceed the minimum acceptable standard of 70% set by Borg and Gall (1983) and 75% set by Fowler (1983), and compare favourably with Mangione's target of 85% for "excellent" response rates (Mangione 1995). They probably reflect the level of patient interest in and commitment to acupuncture and ACUMEN. They also suggest that the motivational strategies employed to minimise loss to follow-up were successful. It is notable that all non-responders at three and six months were in the comparator group, a finding that is suggestive of *"resentful demoralisation"* bias (Torgerson & Sibbald 1998; Campbell & Cook 1979).

6.2.4 Acupuncture treatment protocol

Jonas and colleagues (2002) propose that, *"a treatment being tested must be optimal or at least representative for a defined group of providers and piloted for that condition."* ACUMEN focused on traditional Chinese medical (TCM) acupuncture as practised by members of the British Acupuncture Council. Three practitioners were recruited to this study, so as to improve generalisability and avoid testing the approach of a single practitioner and their preference for points, lifestyle advice and co-interventions (White 2002). The protocol required only that practitioners use acupuncture as the principal therapeutic intervention, and that they provide treatment according to the underlying syndrome diagnosed. However, there were no restrictions on the use of any complementary diagnostic frameworks, such as Five Elements or vital substances, the principal request being that they provide a full and accurate record of treatment in the patient treatment booklets provided. Thus ACUMEN employed a fairly 'open' protocol that allowed the practitioners freedom to use complex and individual approaches for different patients (MacPherson 2004).

Though ACUMEN was designed before the STRICTA recommendations (MacPherson et al 2001), it is reported here in accordance with this consensus document: the style of acupuncture is defined; the practitioners' training, experience and expertise in the treatment of menorrhagia

is stated; and the treatment rationale (including literature sources), regimen, needling details and co-interventions is given.

A more detailed analysis of the treatments reported in ACUMEN is planned by the candidate to study the similarities and differences in TCM diagnoses made within this medically defined population; estimate the mean and range number of points used at a session; and assess the key acu-points prescribed; and the frequency with which co-interventions and advice were used to complement the acupuncture itself. This analysis will consider whether a more precise protocol might be developed for a full trial that would work to facilitate replication without compromising the integrity of the intervention and generalisability. A more tightly defined TCM protocol was reported by Smith and colleagues (2002) for a study of acupuncture to treat nausea and vomiting in early pregnancy, and the authors of the York Acupuncture for Low Back Pain trial (YACBAC) are currently developing a flexible trial protocol with scope for individualised treatment (MacPherson et al. 2004). The qualitative investigation of treatment acceptability to be presented in Chapter Three will guide this process, as will an exploration of the skill base available to a full-scale trial of TCM acupuncture for menorrhagia.

6.2.5 Side effects, reactions and adverse events from treatment

"Feeling drowsy" was experienced by significantly more women in the acupuncture group at three months. This suggests that this common response to treatment is prevalent within this patient group, rather than aggravation of symptoms – a response which is considered more common amongst patients receiving acupuncture for musculoskeletal conditions. Further evidence from a definitive trial would enable health providers to inform women with menorrhagia better about the responses to treatment they might typically expect.

The evidence that the side effect *"stomach pain"* was more of a problem within the usual care group at six months is more difficult to interpret, as the numbers are too small to associate with medication. The effect of acupuncture upon medication use and subsequent differences in side effects and reactions to treatment between groups would be usefully investigated within a full trial. The Medication Change Questionnaire currently being developed by Patterson & Britten (2003) in recognition of the importance of medication avoidance or minimisation in patients accessing CAM might facilitate this.

There were no serious adverse events requiring hospital admission, and just one avoidable event of a forgotten needle. However, there were more reports than expected of minor adverse events in ACUMEN. The majority of patients at three months (67%) and six months (56%) reported pain where the needle was inserted. Whilst pain was one of the most common adverse events reported by patients in a recent national survey, it was described as "prolonged or unacceptable" by only 103 of the 6348 respondents (2%) (MacPherson et al 2004). Therefore it is likely that, given the 100% compliance rates, the majority of patients reporting pain on needle insertion in ACUMEN were referring to a typical and acceptable needle sensation, rather than an event that caused distress and affected treatment acceptability. Indeed this was not an event noted by the practitioners. Similarly, almost a quarter of ACUMEN patients at three months, and almost a third at six months, reported bruising, compared with 0.5% (33/6348) in the national survey. The practitioners reported only four events of bruising in different patients (20%) over the course of the 343 treatments provided in ACUMEN. Slight bruising no larger than the size of a five pence piece and causing no pain or discomfort is common. Thus again the addition of the word "unacceptable" may have made this a more discerning question, and is advised for future trials. Bleeding from the needle site for more than 10 seconds was recorded once by the practitioners (5%) and by five patients (25%), compared with just 0.06% (4/6348) of patients in the national survey spontaneously reporting this event. While a slight drawing of blood is common, bleeding from the needle site for more than 10 seconds is not. Thus these findings suggest either that patients misunderstood this question and recorded any bleeding, encouraged by the tick boxes in the questionnaire, or that the practitioners chose not to record this event. These three questions should be re-worded in future trials to distinguish between acceptable and unacceptable events associated with needle insertion.

6.2.6 Comparator treatment protocol

To assess the effectiveness of normal clinical practice and maximise GP acceptability and participation, ACUMEN chose usual GP care as the comparator intervention. One risk of this strategy was that usual care could have deviated from "best practice" and thus exaggerated effectiveness of acupuncture. In 1999, Prentice wrote that, *"despite widely available evidence inappropriate treatments are being prescribed"*, with only one in 20 GPs prescribing tranexamic acid – *"probably the most effective first line treatment"* – compared with a third prescribing norethisterone – *"arguably the least effective option"* (IMS 1994). As

norethisterone accounted for about 20% of medications prescribed in ACUMEN, a full trial may need to guide the comparator treatment more explicitly.

6.2.7 Choosing outcome measures for menorrhagia

The SF-36 has been validated for this primary care population (Garratt et al 1993). As the Aberdeen Menorrhagia Questionnaire (AMQ) was validated for a secondary care population (Ruta et al 1995), however, a primary care population with less severe symptoms could require a more sensitive instrument to detect clinically significant change. ACUMEN provides some evidence that the AMQ is a valid measure for primary care research, with the scores at baseline ranging from a "mild" (25 out of 100) to "severe" (78 out of 100) GP-assessed menorrhagia score, with a mean of 52. The 351 patients in the development of the AMQ had scores that ranged from three to 90 out of 100, with a mean of 46. The AMQ also proved sensitive to changes in this population, with a significant difference in mean scores detected at six months, despite the maximum sample size of just 39. Future research to establish the psychometric construct validity of this measure for a primary care population would be of value.

The two secondary measures devised to complement the AMQ, by focusing respectively upon three symptoms of specific relevance to menorrhagia (YMRP) and upon 15 common peri-menstrual symptoms (PMSQ), demonstrated sensitivity to change and consistency with the AMQ and SF-36 dimensions of pain and vitality. Further research is planned by the candidate to test the measures' construct validity – that is the extent to which they are related to specified variables in accordance with an established theory or 'hypothetical construct'. For example, the hypotheses to be tested include: those with high scores for pain and heavy bleeding on YMRP and PMSQ will have high scores on the AMQ. The way in which these measures add to our understanding of menstrual health and treatment effects will also be considered further. Full validation of this measure in accordance with the steps prescribed by Streiner & Norman (2001) could be achieved to some extent within a full trial, although a specific study is the ideal. Meanwhile, the candidate plans to compile individual health profiles from the interview data collected at baseline, three and six month follow-up from six of the patients receiving acupuncture in ACUMEN. This analysis will seek to assess the extent to which the health outcomes measured the full range of effects reported by patients, with the aim of informing the

development of questionnaires for a full trial. However, the fact that detailed answers about symptoms was not an objective of the qualitative study may limit the usefulness of findings.

6.2.8 Estimate of effect size from acupuncture and sample size required for a full trial

Statistical power is probably the most crucial concept to address when designing clinical trials (Grimm 2002). It is depends on the current rate of occurrence of the primary outcome in the control group and its expected rate in the experimental group. It also depends on the limits set on errors in the trial: the chance of a Type 1 error (an erroneous positive effect) is typically set at 5%; and the chance of a Type 2 error (a true effect is missed) is typically set at 80%. The findings of ACUMEN suggest that a representative sample of no more than 400 patients would yield robust conclusions.

7 Conclusion

Despite the maximum sample size of only 39, the findings from the ACUMEN trial are most promising, providing consistent, convincing evidence of genuine effects across measures of menstrual and general health with effects evident at 12-month follow-up. Moreover, these gains were reflected in patients reported satisfaction with treatment and treatment outcome, so reducing the risk of an erroneous positive finding (type 1 error). A full-scale trial to evaluate the clinical and economic benefits of offering acupuncture to patients with menorrhagia in primary care is justified. The feasibility of a full trial in light of the experience of ACUMEN is investigated in Chapters Three and Four and discussed further in Chapter Five.

Chapter Four

The ACUMEN Qualitative Investigation:

A longitudinal qualitative study to understand patient perspectives on the provision and acceptability of acupuncture for the treatment of menorrhagia in primary care

1. Background and study rationale

The York and Selby Acupuncture for Menorrhagia (ACUMEN) Exploratory Trial aimed to determine whether future research to evaluate the clinical and economic benefits of offering acupuncture to women with menorrhagia in primary care is both feasible and justified. A key aspect of trial feasibility is patient acceptability, for if women do not consider acupuncture a credible or attractive option, the feasibility of future trials with regard to patient recruitment is uncertain. The justification for investment in future trials will also depend in part upon patient acceptability. For if few women are likely to choose acupuncture, the rationale for evaluating its' provision within National Health Service (NHS) primary care is less compelling.

Past studies have found high levels of satisfaction with complementary and alternative medicine (CAM), both amongst private and NHS users (Luff & Thomas 2000; Gould & MacPherson 2001; Paterson & Britten 1999; Cassidy 1998a; Cassidy 1998b; Kelner & Wellman 1997; Vincent & Furnham 1996). Overall, patients have been found to have a strong, pragmatic, disease-management motivation for CAM use, to develop an increasing commitment to CAM based on positive treatment experiences, and to experience beneficial patient-practitioner relationships. A qualitative study by Luff & Thomas (2000) involving 49 NHS patients, 21 of whom had received acupuncture, also revealed how a commitment to treatment seems to develop as a direct result of user satisfaction with treatment, rather than because of prior belief about complementary therapies, or any commitment arising from financial investments. Moreover, a survey by Gould & MacPherson (2001) involving 72 patients receiving treatment privately from members of the British Acupuncture

Council, suggests that patients can experience a wide-range of benefits in response to acupuncture, that they like the holistic focus of treatment, and are willing to comply with lifestyle advice to aid a return to, or maintain, good health. And in the US survey of patients consulting acupuncturists, the majority reported high levels of satisfaction with their care, of whom 17.4% (100/575) had received acupuncture for gynaecological conditions (Cassidy 1998a & 1998b).

However, how generalisable, or transferable, these findings from heterogeneous patient populations may be to a more homogenous group, namely women with menorrhagia offered the option of referral to acupuncture by their GP as part of a trial, is unclear. For whilst acupuncture is the most popular CAM therapy in the UK, and 55-65% of CAM patients are women, those attending for acupuncture do not tend to seek treatment for gynaecological conditions (Zollman & Vickers 1999). Rather, the most common reason for accessing treatment is a musculoskeletal problem such as pain and stiffness, rheumatoid arthritis and back pain. Indeed, these were found to account for 33% of treated conditions in a national survey by Wadlow & Peringer (1996), whilst gynaecological conditions accounted for 6%, and there is no information about the range of menstrual problems this category included. Thus, acupuncture is not commonly used for gynaecological conditions and patients are therefore unlikely to have heard of it in this capacity, which in turn raises questions about how acceptable they might perceive GP referral to acupuncture for their symptoms of menorrhagia.

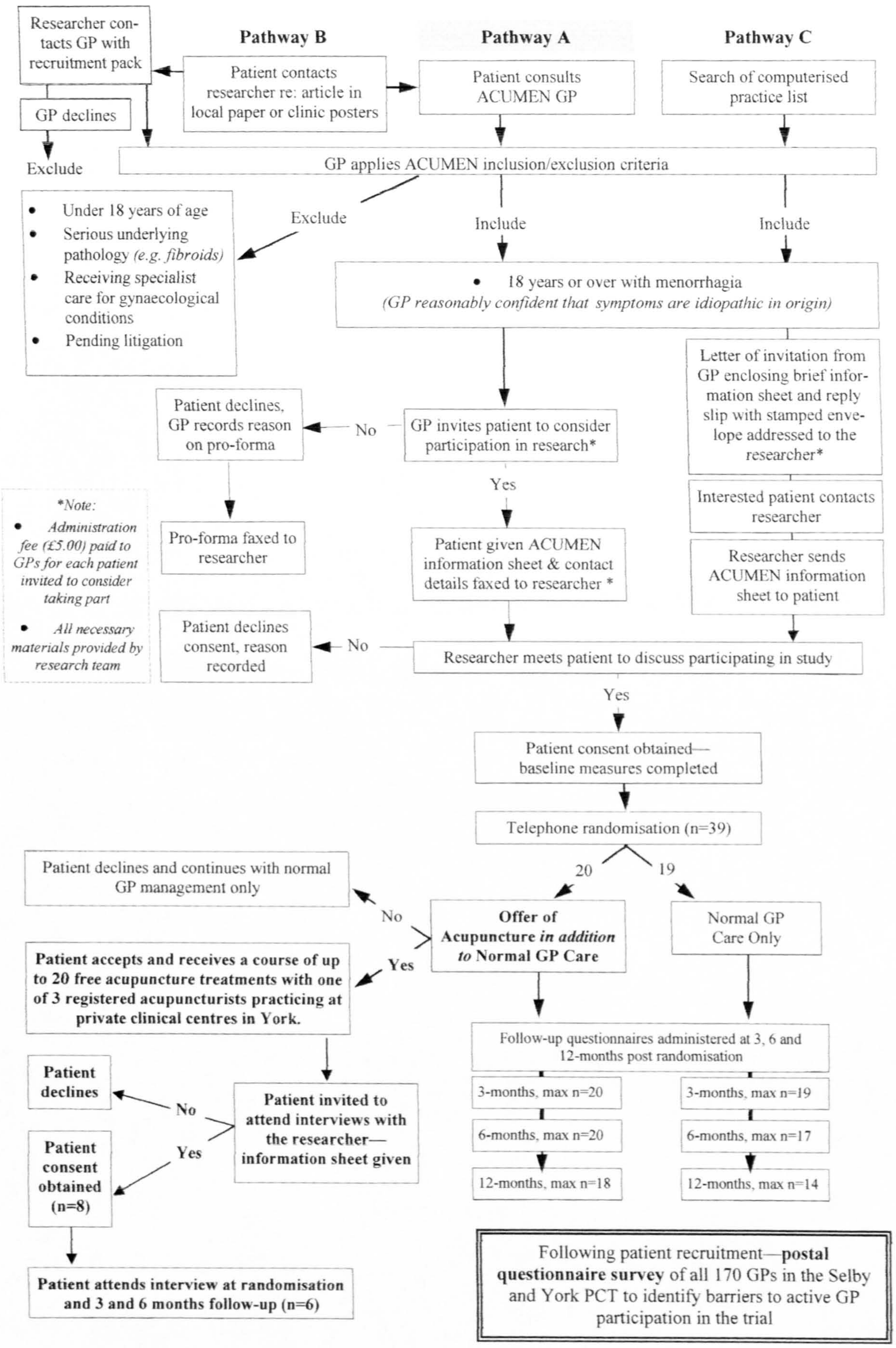
In 2000, Luff and Thomas first identified a need for qualitative studies that purposively examine the experiences of sub-sections of NHS complementary and alternative medicine users, defined by the therapy received and condition treated, and advised that such work be undertaken alongside studies that aim to evaluate the clinical and economic benefits of providing complementary therapies in primary care. Similarly, Cassidy (2001) recommended that qualitative research be used to *"help create clinical trial designs that accurately reflect the wants, values, and needs of research participants,"* as too often, research is designed in accord with the researchers' assumptions, without direct knowledge of the patient population. Whilst Murphy and Dingwall (2001) noted that, *"people act on the basis of what they <u>believe</u> to be true, rather than what might be 'objectively true' (and that) this has important*

implications for the successful implementation of technologies in health, as in other fields, in that advocates for change will need to engage the beliefs of intended users and beneficiaries rather than relying purely on rationalist justifications."

Moreover, it is probable that those factors that first attract patients to use a CAM therapy differ in important ways from the factors that come to determine treatment acceptability as a consequence of lived experience. Indeed, Kelner and Wellman (2000) urge researchers to appreciate health care as a process rather than an event, and argue that whilst retrospective studies are of value they, *"are also limited in their usefulness by the need for respondents to reconstruct their experiences over time"*, and determine that only longitudinal studies can adequately capture this evolution of perception.

Clearly, differences in what first attracted patients to participate in ACUMEN and accept the offer of acupuncture, and the factors that later came to determine treatment acceptability, could have important implications for the design and conduct of any future trials, particularly with respect to recruitment strategies, patient information sheets and delivery of the intervention. Thus a qualitative study that would run parallel to the trial and investigate patient perspectives on the provision of acupuncture for the treatment of menorrhagia in primary care, and in so doing elicit valuable information on the *"complex package of 'active ingredients'"* (MRC 2000) delivered in ACUMEN, was considered integral to the aims of the ACUMEN exploratory study. A longitudinal, prospective design that entailed interviewing a purposive sample of patients at different stages of treatment in the trial (baseline, during and after treatment) was chosen to allow the development of patient perspectives and their relationship to the experience of receiving acupuncture care to be explored *[Figure 7]*.

Fig 7: ACUMEN exploratory trial with nested qualitative study and survey - emergent trial design



2. Aims of the ACUMEN qualitative study

- To explore patient perspectives on the provision and acceptability of acupuncture for the treatment of menorrhagia in primary care
- To inform future research as to the feasibility of and justification for a full-scale trial, given patient perspectives on treatment and the acceptability of GP referral to acupuncture

3. Design

A longitudinal design that entailed interviewing a purposive sample of patients at different stages of treatment in the trial: at baseline, during treatment at three-month follow-up, and after treatment at six-month follow-up.

4. Methods

4.1 Methodological frameworks

This study follows the principles of both qualitative description and framework analysis. Both qualitative description and framework analysis have been identified as methods suitable for applied policy research, where researchers have specific information needs and are intent upon actionable outcomes (Ritchie & Spencer 1994; Sandelowski 2000).

Qualitative description was used to generate informational content relating to acceptability, by first understanding who the women in ACUMEN were in terms of their menstrual history and treatment experiences, and how they came to be involved in the study and why? Framework analysis was used to enable the candidate to go on to provide insights and explanations for whether and why women considered acupuncture an acceptable treatment option for menorrhagia in primary care, given their lived experiences.

4.2 Sample size and selection

The sample frame comprised of those patients randomised to and accepting the offer of acupuncture in the ACUMEN exploratory trial (max=20). Given that this study involved multiple interviews, a sample of six was considered adequate to capture the experiences of women receiving acupuncture in the trial, and eight prudent to allow for either loss of data or loss to interview.

Purposive, reflexive sampling strategies were adopted to maximise the transferability of the study findings (Murphy & Dingwall 2001; Silverman 2000; Miles & Huberman 1994). The initial sampling framework aimed to ensure that interviewees were dispersed amongst the three York-based acupuncture practitioners providing treatment in the trial, and that up to three travelled from outside of York for their treatment. This strategy was amended within the first two-months of recruitment in response to the unplanned recruitment pathways that had emerged. For, whilst the trial protocol had stated that eligible patients would be identified by GPs and invited to consider trial participation during a consultation, it soon became clear that GPs were initiating only a minority of patient referrals [*see Chapters 2 and 4*]. The majority of patient referrals were instead being initiated by the patient, in response to patient-centred recruitment strategies, and facilitated by the GP - the most successful strategies being posters in surgery waiting rooms and a local newspaper feature about the study. Thus the candidate also sought to select patients who had entered the study via different routes *[Figure 7]*.

4.3 The Interviews

The study involved three interviews with each interviewee. All the interviews were conducted by the candidate in the interviewee's own home, and typically lasted between 45 and 90 minutes. Interviews were held at baseline and three and six-months post randomisation, and so coincided with the six-month acupuncture treatment period, and the three and six-month follow-up trial postal questionnaires.

4.3.1 *The interview guides*

The interviews were semi-structured, meaning that whilst the emphasis of the study was on exploration, it was an exploration harnessed by the specific information requirements of the study and "*targeted towards providing 'answers', in the form of greater illumination or understanding,*" for questions about patient perspectives on the acceptability of acupuncture for the treatment of menorrhagia in primary care (Ritchie & Spencer 1994). Thus the candidate used interview guides to shape and direct the interviews (Kvale 1996) *[Table 36 & 37]*.

The principle aim of the first interview was to understand how patients had come to participate in the trial. That is, their motives for joining ACUMEN and expectations about treatment, in the context of their menstrual history and past treatment experiences. Thus, the guide was organised under the three domains: "menstrual history and experiences", "experience of treatments for menorrhagia", and "using acupuncture - motives and expectations". And it was designed to generate answers for the following five research questions:

I. What menstrual experiences do women with menorrhagia have?
II. What are the health concerns of women with menorrhagia?
III. What past experiences of treatment for menorrhagia do women with heavy periods have?
IV. What reasons do women with menorrhagia give for accepting the offer of acupuncture treatment?
V. What expectations as to potential health gains from acupuncture do women with menorrhagia have?

The key aim of the second and third interviews was to understand patient's experiences of acupuncture, and given these, if they considered it an acceptable treatment option for women with menorrhagia in primary care. Thus, this guide was organised under the three domains: "experience of acupuncture", "women's management of their treatment and care", and "acceptability". And it was designed to generate answers to the following six research questions:

I. What experiences of acupuncture do women with menorrhagia have?

II. What are the key components of women's satisfaction or dissatisfaction with treatment?

III. What are the perceived health outcomes from acupuncture experienced by women with menorrhagia?

IV. How do women with menorrhagia who accept acupuncture in addition to standard GP care choose to manage their treatment and care?

V. Do women with heavy periods perceive acupuncture to be an acceptable treatment option?

VI. Do the different components of acceptability unfold chronologically?

These research questions were translated into one or more interview questions, in order that they contribute to a natural conversational flow and thereby generate spontaneous and rich descriptions from the interviewees *[Table 36 & 37]*. The "prompts" were designed to help obtain varied information by encouraging the interviewee to reflect upon their experiences and approach the topic from several angles. However, in order not to impose issues upon interviewees and produce answers to a series of restrictive questions, the candidate did not always ask questions in the exact form or order seen in these guides, but rather asked versions of them as the interviewee's account touched upon the related issues, and viewed the prompt questions in particular as *possible* lines of enquiry (Broom 2005).

The briefing prior to the interview aimed to put the interviewee at ease, whilst telling them the purpose of the interview and how it would be structured. It also re-affirmed their willingness for the interview to be recorded, and invited them to ask questions before starting the interview proper. The de-briefing at the end of each interview provided an additional opportunity for issues the interviewee had been thinking or worrying about during the interview to be dealt with, either on or off tape.

Prior to each follow-up interview, the candidate read the transcripts from the previous interview and took note of the accompanying comments she'd recorded in her field diary about the key themes or issues that had arisen at the previous interview and suggested lines of future enquiry.

Table 36: Baseline interview guide

Briefing: *"What I'd like to talk about with you today is the history of your periods and the treatments you have had. I'd also like to ask you about your thoughts about acupuncture treatment. Does this sound okay to you? Is the tape recorder okay there, do you think?"*	
Research Questions	**The 3 domains and interview questions**
1. What menstrual experiences do women with menorrhagia have? 2. What are the health concerns of women with menorrhagia?	***A. Menstrual History and Experiences*** 1. To begin with, could you tell me the brief history of your periods, from when they started up to now? *Prompt:* ➢ *Have your periods always been the same?* ➢ *Do you have any ideas about what might have caused or triggered them?* 2. In what ways, if at all, have your periods affected or interfered with your everyday life over the years? *Prompt:* ➢ *How does having heavy periods make you feel?* 3. What concerns you most about your periods?
3. What past experiences of treatment for menorrhagia do women with heavy periods have?	*B. Experience of Treatments for Menorrhagia* 1. What treatments have you tried over the years to help with your periods? *Prompt:* ➢ *What treatments has your GP prescribed for you?* ➢ *Have you tried anything else at all?* ➢ *What are your thoughts on treatment x?* ➢ *What would you say are the pros and cons of treatment x?* ➢ *What usually helps you feel better/worse during a period?*
4. What reasons do women with menorrhagia give for accepting the offer of acupuncture treatment? *5. What expectations as to potential health gains from acupuncture do women with menorrhagia have?*	*C. Using acupuncture - motives and Expectations* *1. Can you tell me about how you first heard about the ACUMEN trial, and why you accepted the offer of acupuncture?* *Prompt:* ➢ *What do you expect acupuncture to be like?* ➢ *What ideas do you have about what treatment will involve?* ➢ *What do you expect your acupuncturist to be like?* ➢ *Do you think acupuncture can help?* ➢ *Ideally, what do you hope treatment will do for you?* ➢ *What do you expect from treatment?* ➢ *What ideas do you have about how it works?*
De-briefing: *"That's great, thank you so much for your time. I have no more questions, but is there anything else you think I should know before we finish? Or maybe something that you'd like to ask me about? Shall I turn the tape off?"*	

Table 37: Interview guide at three and six-month follow-up

Briefing at three-month follow-up: *"It's been 3 months now since you joined the study and started your course of acupuncture, so what I'd like to do today is find out about the next part of the story. You know, talk about how you've been since we last met, and your thoughts and feelings about acupuncture, and any other treatments you've tried. Does that sound okay?" Is the tape recorder okay there, do you think?"*
Briefing at six-month follow-up: *"It's been 6 months now since you joined the study and your acupuncture has probably finished now, or is about to finish, and this is our last interview, so what I'd like to do today is find out the final chapter in the story. You know, talk about how you've been since we last met, how you've found having acupuncture these past 6 months, and also where you think you'll go from here, in terms of treatment for your periods? Does that sound okay?" Is the tape recorder okay there, do you think?"*

Research Questions	The three domains and interview questions
6. What experiences of acupuncture do women with menorrhagia have? 7. What are the key components of women's satisfaction or dissatisfaction with treatment? 8. What are the perceived health outcomes from acupuncture experienced by women with menorrhagia?	***A. Experience of Acupuncture*** 1. To begin with, could you tell me about your acupuncture treatment? What has it been like and how effective do you think it has been? Prompt: ➢ Was it what you expected? ➢ How do you feel when you're having acupuncture? ➢ What has your acupuncturist explained to you about your treatment? ➢ Over the past 3/6 months, would you say your periods have become better, worse, or are the same as they were before you joined the study? ➢ *What sort of changes have you been aware of?* ➢ *What sort of change do you expect, or are you hoping for, over the next 3 months?* ➢ *What sort of changes had you expected, or hoped for, over the past 6 months?*
9. How do women with menorrhagia, who accept acupuncture in addition to standard GP care, choose to manage their treatment and care?	***B. Women's management of their treatment and care*** 2. I'm also really interested in what else you've tried, alongside the acupuncture, to help with your periods. Could you tell me about any other treatments or remedies you've tried? *Prompt:* ➢ *What treatments has your GP prescribed for you?* ➢ *How did you get on with x?* ➢ *What did you hope to get from x?* ➢ *Have you tried anything else at all, supplements or?*
10. Do women with heavy periods perceive acupuncture to be an acceptable treatment option? 11. Do the different components of acceptability unfold chronologically?	***C. Acceptability*** 1. Knowing what you know now about acupuncture, if you were able to travel back in time, do you think you would still make that same choice to have acupuncture, or would you choose to try something else? 2. What advice would you give to other women who might be thinking about whether or not to try acupuncture for their heavy periods? *<u>At 6-month follow-up choose between 3a or b as applicable:</u>* 3a. If over the coming months or years your periods become heavy again, what treatment would you want? 3b. If tomorrow you were offered more acupuncture to help with your heavy periods, do you think you'd accept it? *Prompt:* ➢ *Would you like to try a different treatment now?* ➢ *What do you think your next step will be, with regards to finding an effective treatment for your heavy periods?*

De-briefing: *"That's great, thank you so much for your time. I've no more questions, but is there anything else you think I should know before we finish? Or maybe something that you'd like to ask me about? Shall I turn the tape off?"*

4.4 Research governance

4.4.1 Informed consent

At randomisation, the candidate provided all patients fulfilling the selection criterion with an integral information sheet and consent form *[Appendix 38]*. This explaincd the purpose of the study, when and how the interviews would be conducted, and the data management procedures in place to address issues of confidentiality and anonymity *[see 4.5]*.

To encourage as uncensored an account as possible and to avoid the potential for response bias, the interviewees were not informed of the candidate's professional status as an acupuncturist. Instead, interviewees were allowed to assume the candidate's role did not extend beyond that of a student in health sciences.

4.4.2 Ethical approval

The candidate, with the support of her supervisory team, first applied for ethical approval for the ACUMEN exploratory trial and nested qualitative study in January 2001. Approval was granted in March 2001, once the ACUMEN research team had successfully addressed all those issues raised by the committee in response to the initial application. A copy of the letter accompanying the revised application and subsequent letter of approval received from the Chair are given in Appendices 31 and 32.

4.4.3 Data management

The collection and management of data was carried out in line with the Data Protection Act 1998, and in accordance with the Policy for the Protection of Data and Data Providers prescribed by the Department of Health Sciences, University of York. The interviewees' contact details were entered into the master index at the research centre by the candidate and password protected. All other data held either on computer, or in a locked filling cabinet at the research centre, was coded with the interviewees unique study number and pseudonym. This included the tapes, transcripts, and field diary entries. Anonymity was achieved during transcription by

replacing all names of individuals or places with identifiers such as 'daughter', 'doctor', 'acupuncturist', 's/he' and 'clinic.' Interview tapes are kept under lock and key and will be erased following completion of the research.

4.5 Data analysis

4.5.1 Transcription

All the interviews were fully transcribed by the candidate. In most instances, transcription took place soon after the interview in order to aid the inclusion of bodily expressions accompanying statements, as these often provide richer access to meanings (Kvale 1996). Where this was not possible, the tapes were listened to and notes made about the relevant segments of speech and accompanying bodily expressions.

4.5.2 Qualitative content analysis (descriptive)

The data from the initial interview was analysed according to the principles of qualitative description using qualitative content analysis, which is orientated toward summarising the informational contents of the data (Sandelowski 2000). The analysis began by using the three domains from the interview guide as a coding system: "menstrual history and experiences", "experience of treatments for menorrhagia", and "using acupuncture - motives and expectations". This coding system was modified over the course of the analysis to separate women's motives for using acupuncture from their expectations of acupuncture.

The descriptive analysis of the initial interviews led first to the creation of case charts where verbatim text was regrouped according to the coding scheme - a simple descriptor and reference was used to 'tag' some chunks of text. The aim here was to avoid a distilling down and loss of data and instead organise this rich textual data to better facilitate access to the patients' own words. Verbatim, or tagged text for each code was organised under a distilled summary. The original text was referenced according to interview and line number so that the source could be traced back and the process of descriptive analysis examined and replicated. These case charts facilitated first, the case-based summaries in section 5.2 below that attest to the uniqueness of

the cases, and second, a cross-case account of the interviewees' shared and divergent menstrual experiences, motives for joining ACUMEN and expectations of acupuncture care. The case charts are identified by the interviewees' research name (Linda, Sarah, Ann, Clare, Julia and Lyn) and included as Appendices *[Appendices 39 to 44]*. The quotes were chosen on the basis of how well they illustrated the themes identified and the range of views or experiences reported.

4.5.3 Thematic analysis (interpretive)

The data from the follow-up interviews was analysed according to the principles of framework analysis, as set out by Ritchie and Spencer (1994). This analysis enabled the candidate to generate a thematically organised account of the meaning of treatment acceptability to the women participating in interviewees and receiving acupuncture for their symptoms of menorrhagia in primary care for the purposes of the ACUMEN exploratory trial. This analysis involved five key steps or stages:

I. Familiarisation,
II. Identifying a thematic framework,
III. Indexing or coding,
IV. Charting,
V. Mapping and interpretation.

The purpose of the familiarisation stage is to assess the richness, depth and diversity of the data, and in so doing, begin the process of abstraction and conceptualisation. Thus the candidate became immersed in the totality of the data through listening to the tapes, reading the transcripts and studying observational notes made in her field diary, and at the same time noted key ideas and recurrent themes related to treatment acceptability. These notes then facilitated the identification of the thematic framework within which the material was sifted and sorted. In studies involving the analysis of a large number of transcripts (20 or more), the thematic framework would be developed using a selection of the transcripts and once refined, it would be systematically applied to all the transcripts in the process of indexing or coding. In this study, because of the relatively small number of transcripts (six three-month follow-up transcripts and six six-month follow-up transcripts) the development and refinement of the framework, or index, took place at the same time as coding the

transcripts, and involved a to-ing and fro-ing between the coded transcripts and index in order to continually test assertions made about the relevance of themes and the connections between ideas. This analysis led to the identification of five themes related to treatment acceptability in ACUMEN:

I. A safe, natural alternative to drugs and surgery
II. The acupuncture treatment process
III. A holistic approach to treatment and care
IV. Perceived health outcomes from acupuncture
V. Patterns of use of conventional medicine and acupuncture

Some of these themes were informed by the research aims and were introduced into the interviews via the interview schedule (experience of the acupuncture treatment process, perceived health outcomes from acupuncture, use of conventional medicine and acupuncture), whilst others were emergent themes that came from issues raised by the respondents themselves (a safe, natural alternative to drugs and surgery, a holistic approach to treatment and care). The accompanying sub-themes were developed in the same way as the themes, and worked to encapsulate the key issues within each theme. An 'other' column was included for data that did not neatly fit this schema and worked to ensure that information was not lost in the process of analysis *[Table 38]*. And whilst the themes were not verified independently, the emergent framework was agreed by the candidate's supervisor (KT) (Silverman 2000; Miles & Huberman 1994).

Having coded the data in accordance with the refined framework, chunks of verbatim text were then 'lifted' from their original context and rearranged according to the appropriate thematic reference to generate thematic charts *[Appendices 45 to 49]*. This lifting of verbatim text to create each chart is contrary to the abstraction and synthesis advised by Ritchie and Spencer (1994), where a distilled summary of each respondent's views is entered on the chart. The reason for this approach was to avoid a distilling down and loss of data and instead organise this rich textual data into manageable 'bites' within the thematic framework so as to better facilitate access to the patients' own words. Organising the entire data set in this way, also allowed the candidate to test the reliability of findings, by looking continually for disconfirming evidence. Verbatim, or tagged, text was referenced according to interview and line

Table 38: A thematic framework to facilitate an exploration of the meaning of treatment acceptability for women participating in interviews and receiving acupuncture for menorrhagia in primary care for the purposes of ACUMEN

I. CHART ONE: A safe, natural alternative to drugs and surgery a. "It's more natural isn't it"
II. CHART TWO: The acupuncture treatment process a. "Just a little pin prick" b. "You feel so relaxed" c. "It does take time" d. "Someone listening, someone understanding" e. Other (e.g. massage)
III. CHART THREE: A holistic perspective a. "Feeling better completely, as a whole" b. "The diet thing as well, that's been really good" c. Other
IV. CHART FOUR: Perceived health outcomes from acupuncture treatment a. "So easy now, so light", "it's manageable now, you know, I can cope", "heavy, heavy enough to worry me" b. "The treatment I had had an affect on all of me" c. "I'm making a conscious effort to keep relaxed and keep stress free and keep healthy really" d. Other (e.g. Mirena coil fitted)
V. CHART FIVE: Acupuncture and conventional medicine: patterns of use a. Perceived role for acupuncture and conventional medicine b. Predicted future use of acupuncture

number so that the source could be traced back and the process of thematic analysis examined and replicated. Each chart theme or sub-theme was labelled iteratively by paraphrasing interviewees' words. A summary of the diverse range of views and experiences was also included where it might aid recall and facilitate analysis *[for example, see the column "other", chart 2: the acupuncture treatment process, Appendix 46]*.

Once all the data had been sifted and charted according to core themes, the candidate began to map and interpret the data set as a whole. She reviewed the charts, comparing and contrasting the interviewee's perceptions and experiences of referral to acupuncture as part of a trial, and looked for patterns and connections by testing and developing the ideas that had first begun to emerge during the familiarisation stage of the analysis. The quotes given, as with the descriptive analysis, were chosen on the basis of how well they illustrated the themes identified and the range of views or experiences reported.

4.6 Outputs

Three different outputs were obtained. These involved either the data collected at the initial baseline interview, or the data collected from the two follow-up interviews held during and after each interviewee's course of acupuncture treatment. These are described, along with the source material for the analysis, in Table 39 below:

Table 39: Study outputs and source material

	Output	Source material
I	Demographic characteristics of the sample and acupuncture group	Trial questionnaire data
II	Case-based summaries of the clinical history of each patient as given by them in the interview and the events that led them to participate in ACUMEN	Data from the initial, baseline interview (six transcripts)
III	Cross-case, descriptive account of the sampled patients' shared and divergent menstrual experiences, motives for joining ACUMEN and expectations of acupuncture care	Data from the initial, baseline interview (six transcripts)
IV	A thematically organised account of the meaning of 'treatment acceptability' to the women participating in interviews and receiving acupuncture for menorrhagia in primary care for the purposes of ACUMEN.	Data from interviews held during and post treatment (12 transcripts) with reference to the descriptive analysis of the initial interview data

5. Results

5.1 The sample

All eight of the patients invited gave their informed consent to take part in semi-structured interviews with the candidate. Two of the interviewees were excluded from the final analysis: one because they'd been unable to attend a third interview; and one because the recorded data from their third interview proved untranscribable. Thus this study involved the analysis of interviews with six of the 20 women randomised to and accepting the option of acupuncture in the ACUMEN trial. Their demographic characteristics, by way of entry route to the study, acupuncture practitioner, age, education, housing and employment are given in Table 40, alongside the characteristics of the acupuncture group as a whole. The similarities and differences between those in the sample frame and those interviewed are described below.

5.1.1 Demographic characteristics of the sample and acupuncture group

The demographic characteristics of the sample and acupuncture group as a whole were generally similar. The sample comprised married, white, English women aged between 34 and 49. Most had one or two children and lived in semi-detached or terraced housing that they owned with a mortgage. The majority lived in York as opposed to Selby, and most travelled to their acupuncture treatments by car. The sample covered the full range, and reflected equally the views of women who had left education at around 16, 18 or 20-years of age. However, the sample differed to the sample population (acupuncture group) on the employment dimension, as women engaged in interviews were more likely to be full-time looking after the home and family and less likely to be employed full-time.

The sample and acupuncture group were equivalent in terms of what led to their choice of ACUMEN acupuncture practitioner, with the sample including women who'd been guided by their preference for a local clinic, a female practitioner, or to remain in the care of their own practitioner. The two dimensions the sample and acupuncture group differed on were entry route into the study and dispersion amongst

Table 40: Characteristics of the acupuncture group and the six interviewees

	Range for acupuncture group n=20	**Linda**	**Lyn**	**Sarah**	**Ann**	**Clare**	**Julia**
Trial entry route	Surgery poster = 8 Acupuncturist = 5 GP referral = 2 Newspaper feature =3 Letters to a GPs eligible patients = 2	Surgery poster	GP referral	Acupunctu rist	Surgery poster	GP referral	Surgery poster
Acupuncturist & no. patients	A = 9 patients B = 5 patients C = 6 patients	A	B	C	B	C	B
Choice of practitioner	Most local = 8 Prefer female = 6 Own practitioner = 4 Clinic times = 2	Most local	Prefer female	Continue with own practitioner	Prefer female	Most local	Prefer female
Home location	York = 14 Selby = 6	York	Selby	York	Selby	York	York
Travel to acupuncture	On foot =3 Bike/scooter=2 Car =12 Bus =2 Missing =1	Car	Car	Car	Bus	On foot	Bike
Age	18-29yrs = 1 30-39yrs = 5 40-49yrs=13 50+ = 1	44	42	34	49	37	38
Marital status	Single = 1 Married = 16 Divorced= 3	Married	Married	Married	Married	Married	Married
Number of children	0 = 1 1 to 2 = 14 3 to 4 = 5	1	3	1	4	2	2
Age on leaving education	16 or less=10 17 to 19 = 7 20+ = 3	16 or less	17 to 19	20+	16 or less	20+	17 to 19
Higher education	Yes = 16 No = 4	No	Yes	No	No	Yes	Yes
Type of home	Detached house =2 Semi-detached =8 Terraced =8 Flat=1 Maisonette=1	Semi-detached house	Terraced house	Detached house	Terraced house	Semi-detached house	Terraced house
Owns own home?	Outright=3 Mortgage =11 Rents=5 Lives here rent free=1	Mortgage	Mortgage	Mortgage	Lives here rent free	Mortgage	Mortgage
Employment	Looking after home and family = 5 Full-time = 5 Part-time = 8 Unable to work for poor health = 2	Part-time	Full-time	Part-time	Looking after home and family	Looking after home and family	Looking after home and family
Ethnicity	White English = 19 White Other= 1	White English	White English	White English	White English	White English	White English

the three ACUMEN acupuncture practitioners. This occurred first because the women sampled had entered the trial via the three routes in operation at the time of recruitment to interview. Namely, in response to a poster in the GP surgery, on the

recommendation of their acupuncturist, or at the invitation of their ACUMEN GP, and it is worth noting that the surgery poster and acupuncturist recommendation were the most common trial entry routes within the acupuncture group. However, the decision to test other patient-centred recruitment strategies after recruitment to interview had been completed meant that the study did not include the views and experiences of those patients who had responded to a feature about the study in a local newspaper, or to a letter of invitation from their GP. And second, whilst the dispersion amongst ACUMEN acupuncture practitioners had been relatively even at the time of recruitment (practitioner A, three: practitioner B, three: practitioner C, two), after the two losses to follow-up this dispersion became uneven (practitioner A, one: practitioner B, three: practitioner C, two), and because the losses were both patients of practitioner A, the dispersion also does not reflect the fact that practitioner A had treated almost half the acupuncture group on account of their working more clinic hours than the other two practitioners.

5.2 Case-based summaries

The following six case-based summaries describe the clinical history of each interviewee, and the events that led them to participate in ACUMEN - as given by them in the initial, baseline interview. These summaries aim to establish the uniqueness of the cases, before the distinctiveness of their personalities is set-aside in a cross-case analysis of their shared and divergent experiences and perspectives on ACUMEN *[5.3]*. Each case summary is organised with reference to the coding system: menstrual history and experiences; experience of drug therapies; using acupuncture – motives; and using acupuncture – expectations; and identified by the research name given to each interviewee:

I. *Linda [5.2.1]*,
II. Lyn *[5.2.2]*,
III. Sarah *[5.2.3]*,
IV. Ann *[5.2.4]*,
V. Clare *[5.2.5]*, and
VI. Julia *[5.2.6]*.

5.2.1 Linda

"I get big clots that come away from me...and I always think that maybes its all your childbearing things coming away from you. / It makes me very tired, I just get so very tired. / ...(and) I honestly didn't think that there was anything that they (GPs) could do."(1:43-8; 108; 138-43)

Linda was 44-years of age, married with one teenage daughter and employed part-time as a support worker for adults with learning disabilities. She had had irregular, painful and heavy periods since menarche at 15 years of age, which had become increasingly heavy and painful over the past 10 years, with flooding[1], the passing of large clots, pain travelling down her legs and achy finger and wrist joints. Linda had wondered whether the worsening of her symptoms was evidence of a natural deterioration of the womb and prelude to the menopause. Her primary health concern was anaemia, and she consulted her GP to ask for an iron supplement whenever she felt tired and worn down. Linda had found these effective at boosting energy levels, but used them with caution, as they also caused constipation. She had never asked specifically for treatment for heavy periods, as she did not believe GPs had anything to offer.

It was an ACUMEN poster in the doctor's surgery that led to Linda's participating in the trial. Her key motive was to find effective treatment, but also expressed concerns about side effects from both drug therapy and iron supplements, and a preference for natural, low-tech interventions. Therefore her wish to use acupuncture was also closely linked to understanding it as a safe, natural, low-tech therapy that stimulates the body's own healing resources to bring about health. She was reassured by acupuncture's long history, along with positive media reports of acupuncture in other contexts (back pain, anaesthesia), and her asserted rationale that other women with menorrhagia must have been found to benefit from treatment for it now to be being used in a trial. Linda expected treatment to feel similar to the acupuncture she'd received for her shoulder from her chiropractor, and given this past experience and the

[1] Linda did not mention flooding at her initial interview, although she described it as a serious problem at *follow-up interviews – see Appendix 13, chart 4: outcomes.*

adverse events listed in the ACUMEN information sheet, expected some bleeding and bruising from the needles *[for case chart see Appendix 37].*

5.2.2 Lyn

It's a wearing down process isn't it, if you have heavy periods every month? And I tend to think, "oh no, not again." / (But) generally I just think it's something I have to put up with, because I don't really want to have a hysterectomy. You know, I do feel it is there for a purpose - my womb. " (1:250-6;178-88; 192-201; 241-8)

Lyn was 42-years of age, married with three teenage children, and employed full-time as a teaching assistant. Her periods had not been a cause for concern until after the birth of her third child and discontinuing the oral contraceptive pill (OCP), when they had become increasingly heavy with sudden flooding and severe pain that had caused fainting. The pain tended to pre-empt passing large clots, and she usually woke twice during the night to change. Prior to the bleed, Lyn experienced irritability, tiredness, mood swings, poor appetite and craving chocolate. Post bleed, Lyn felt washed-out and light-headed. Her cycle had become shorter (three-weekly) over the previous year, and she was concerned about the long-term effects of heavy blood loss upon her health. She was also fed-up with coping with her menstrual symptoms.

Lyn had consulted her GP for treatment over the previous five years, but the drug-therapies prescribed had become increasingly ineffective over time (OCP, tranexamic acid and others). Therefore, at her last annual review, Lyn's GP had suggested she consider participating in the ACUMEN trial. Lyn had agreed, as she liked that acupuncture might provide her with an effective alternative to both drug therapy and surgery. She also expressed concerns about side effects, and stated that whilst hysterectomy often seemed to be a solution, she did not consider it an acceptable option. Other factors that influenced her decision were her GP's positive attitude to the trial, positive reports of acupuncture used in other contexts (dieting and stopping smoking), and her own belief it would be of benefit. And whilst she was unsure how acupuncture could bring about health and healing, she expected treatment to affect all aspects of her health, addressing both physical and psychological symptoms to bring about a sense of well-being. She also expected acupuncture to involve the insertion of

needles at specific sites on the body, and for her practitioner to wear a white uniform *[for case chart see Appendix 38].*

5.2.3 Sarah

"Your body is tired, your mind's tired, you lose your concentration, you're not that great at your job because your mind it's, you know, you're just so tired.../... It has affected every part of my life./ ...(and) I don't want anything affecting my fertility side of it, because if, you know, I stop ovulating that would be even worse..." (1:181-3; 548; 294-304)

Sarah was 34-years of age, married with one child under five, and employed part-time as a hotel manager. She had experienced amenorrhoea in her teens, but after discontinuing the OCP in her early-twenties, Sarah's menses had become regular, with a light bleed and mild abdominal discomfort - a pattern that had continued after the birth of her son. During the 18-months prior to joining the study, Sarah had experienced a miscarriage followed by an ectopic pregnancy and a further miscarriage, all within the first trimester. The second period following the last miscarriage was uncharacteristically heavy, painful and prolonged with mood swings. The bleeding had become lighter, but had continued with little respite for three-months. She had found the tiredness from the continual bleeding "devastating", and was acutely aware of its impact upon her ability to conceive. She had consulted the GP's at her local surgery regularly because of concerns about serious underlying pathology, but was disappointed by their lack of empathy and inability to prescribe effective treatment - an increased risk of developing thrombosis and her concerns about disrupting ovulation excluded tranexamic acid and hormone treatments.

Sarah's decision to use acupuncture had been guided by positive published reports - including an article in The Times newspaper, her niece's experience of acupuncture, and acupuncture's long history. She understood acupuncture to be a safe, natural and holistic therapy that would treat the underlying cause, and was particularly effective at calming the mind to regulate the function of the body and make conception possible. Indeed, Sarah had expressed clearly defined beliefs as to the limitations and strengths of both conventional medicine and acupuncture, and was able therefore, to predict future patterns of use: namely that she would use acupuncture for symptoms arising

from stress (e.g. hormone irregularities and prolonged menstrual loss, indigestion and lack of energy) and conventional medicine to diagnose and treat disorders arising from serious underlying pathology (e.g. ectopic pregnancy and thromboembolic disease).

Following her first treatment, Sarah had been impressed and reassured by her acupuncturist's counsel and prognosis and the needle effects. The holistic perspective taken had also validated her experience of ill health, and provided a conceptual framework by which to understand the ways in which different aspects of her physical and mental health were connected and had been affected by her heavy, prolonged bleed. It was Sarah's acupuncturist who had prompted her to request a referral to the study from her GP, and reassured her that prolonged as well as heavy menstrual loss met the inclusion criteria *[for case chart see Appendix 39]*.

5.2.4 Ann

"You think, "oh, am I ever going to make this up again?" You know? "No wonder I feel tired, no wonder I feel weak"." / (The tranexamic acid) maybe would work if I could take them properly...(but) I don't like taking tablets, quite honestly...(and with) things like hysterectomy you're out of action for quite a few weeks." (1: 110-112; 123-155; 207-216)

Ann was 49-years of age, married with four children aged 10-years and over. She cared for them and her terminally ill husband at home. Since her menarche at 16 years of age, Ann had had heavy, painful, tiring periods, with the bleed lasting eight to 10 days. The bleed worsened when an intra-uterine device (IUD) was fitted and did not improve following sterilisation. She woke during the night to change, and more recently had experienced pains running through her hips and down her legs, and debilitating migraine headaches. Ann was concerned about the long-term effects of heavy blood loss and was looking forward to her menopause. Over the years, Ann had received unsuccessful, or unacceptable, GP treatment for her period symptoms that had included the oral contraceptive pill (OCP), an IUD, sterilisation, hormone replacement therapy (HRT) and most recently, tranexamic acid. Ann acknowledged that her concerns about side effects from the drug-therapies prescribed had led to poor compliance and to her relying upon over-the-counter remedies such as Evening

Primrose Oil and iron supplements. She had gained only limited benefits from these, and no lessening in menstrual flow or pain, but continued to take an iron supplement occasionally to help prevent anaemia. She asserted that she considered hysterectomy to be an unacceptable treatment possibility, given her family commitments and the recovery time that would be required.

Ann had enquired about ACUMEN in response to leaflets displayed in the GP surgery. She was pleased with her GP's positive response to her enquiry and delighted that she'd had to wait just a few days before hearing from the researcher (candidate). Her own positive experience of acupuncture for neck pain, along with positive reports from others who had used acupuncture, had influenced her decision to participate. Ann understood acupuncture to be a safe, natural, low-tech intervention, and liked that attending for weekly acupuncture would provide regular time for self. She had expected treatment to be similar to the acupuncture she'd received from her GP for her neck, which involved needles being inserted locally, retained for 20 minutes and manipulated every ten. Ann expressed some anxiety about the needles, but was prepared to over-ride these for the expected benefits *[for case chart see Appendix 40]*.

5.2.5 Clare

"...I've ruined my car seat...it just makes you feel horrible. / And the other side of it is the sex side of it...I feel so sore...you don't feel nice... / (And) is the pain an indication of something worse going on down there that I don't know about?...((Cries)) / And I always feel as though I'm fighting for energy. And I know you get more tired as you get older, but I am only 37, I feel as though I aught to be a bit bouncier than I am. But I'm not ill, but I'm not right either." ((Cries)) (1:152-159; 237-244;222-236; 251-256)

Clare was 37-years of age and married with two children aged four and six years. By talking to friends, Clare had come to realise that she had had heavy, painful periods since menarche at 13-years of age. From 16 to 28-years of age, the oral contraceptive pill (OCP) and prescribed non-steroidal anti-inflammatory drugs (NSAIDs) lessened the pain, the bleed and addressed the acne - although the pain still caused absenteeism. Clare had described her periods as "manageable" after discontinuing the

OCP to have her children, but they had worsened considerably after a difficult vaginal birth with her second child four years previous. Since this time, her periods had been heavier with sudden flooding, passing large clots and severe pain. She had also experienced acute pain with bowel movements and during intercourse, a dragging period-like pain with walking, and fatigue. Clare had been concerned that her symptoms were a sign of serious pathology and was troubled that four years of largely ineffective treatment had not alarmed her GPs. She had been distressed that a natural female phenomenon should feel like an illness, and her symptoms so adversely impact upon her quality of life and sense of self.

Clare's GP had suggested referral to ACUMEN, and a local clinic that would make attending for acupuncture possible in terms of time, was an important factor in her decision to participate. Clare liked the idea of a non-drug therapy that would work to bring about, or restore, normal physical functioning, and felt that this might be just what her body needed to recover fully from the difficult birth of her second child, but expected treatment to be ineffective for symptoms arising from organic pathology, such as pain from scar tissue. She understood acupuncture to involve the insertion of needles at specific sites along nerve or energy lines, and that points that may not be near the uterus to be connected to and have an affect on the uterus, in a similar way to how referred pain responds to treatment distal to the site of pain. Clare expected her acupuncturist to be similar to her doctor, and for there to be some discomfort from the needles, given her husbands painful experience, but this expectation was tempered by her understanding that acupuncture doesn't always hurt *[for case chart see Appendix 41]*.

5.2.6 Julia

"I have to be careful if I do go out, I mean I can go out, except in summertime when you don't wear such heavy clothing .... / My concern is not knowing why...It's the not knowing why it's so heavy. Why am I suffering like this? I want to know the reasons...I want a cure...I mean I talk about it like it's an illness, which is, you know, but it does feel like that sometimes..." (1: 214-9; 244-53)

Julia was 38-years of age and married with two children aged seven years and over. She'd had painful periods since menarche, and at 16-years of age had been prescribed the oral contraceptive pill (OCP), which successfully addressed the pain. Julia discontinued the OCP after the birth of her second child six years ago, due to concerns about long-term usage. Her periods had since been very heavy and painful, with sudden flooding and passing large clots. Julia had also experienced pain at ovulation, with bloating, tiredness and poor concentration prior to and during her period, which had a considerable impact upon her quality of life. She showered twice daily during her menses, avoided outings and needed to sleep in the day. Not knowing why a natural process felt like an illness had caused her anxiety, and she was concerned about the long-term effects of heavy blood loss upon her general health. Over the previous six years, Julia had received unsuccessful treatment for pain, heavy bleeding and bloating, and now used over-the-counter painkillers and vitamins. She had been prompted to consult her GP again by the ACUMEN poster in the surgery waiting room, and had been pleased with the empathy her GP demonstrated at this consultation, referral to the study, and the prescription for tranexamic acid. This sympathetic reception had also raised her interest in returning to enquire about hormone tests.

Julia's decision to join the study had been motivated by her interest in CAM and concerns about side effects from drug therapy. She had also hoped that the study might provide her with an explanation for her symptoms. Julia had been encouraged by positive reports from people she knew who had used acupuncture, but did not have high expectations. She believed acupuncture to involve the insertion of needles at specific sites - possibly related to nerve endings, and in a similar way to reflexology, expected acupuncture points that may not be near the uterus to be connected to and have an affect on the uterus *[for case chart see Appendix 42]*.

5.3 Shared and divergent menstrual experiences, motives for joining ACUMEN and expectations of acupuncture care

The case-based summaries above establish the uniqueness of the cases. The purpose of the following cross-case, descriptive account is to setout the sampled patients' shared and divergent menstrual experiences, motives for joining ACUMEN and expectations of acupuncture care. This account is organised with reference to the coding system and concludes with a summary of the findings. The headings under which findings are organised are as follows:

VII. Current symptoms *[5.3.1]*,
VIII. Symptom history *[5.3.2]*,
IX. Previous treatment experience *[5.3.3]*,
X. Motives for joining the trial *[5.3.4]*,
XI. Expectations of acupuncture care *[5.3.5]*, and
XII. Summary *[5.3.6]*.

5.3.1 Current symptoms

The interviews held at baseline revealed shared symptomatology amongst the interviewees. All had experienced an excessive or prolonged menstrual bleed in conjunction with pain at the time of the period (dysmenorrhoea), and tiredness. The other most commonly associated symptoms were the passing of large clots and sudden flooding. Three interviewees reported clots and flooding at the first interview and one mentioned it only in her later interviews, despite the fact that she describes it as a serious problem *[see 5.2.1 footnote]*.

The symptoms experienced at menstruation had adversely affected the interviewees' quality of life. They had all described how their ability to carry out daily tasks had been affected, and two had told how it had affected their relationship with their husband, particularly with regard to sexual intimacy:

> *"I daren't walk too far. Obviously I have to be very heavily padded up so I don't like going out far. If I have to go out I go, but I don't like going far. I try and avoid it. (Ann, 1:228-45)*

"And you feel then that there's not much time within a month where you do actually feel good, and feel yourself. / So when I do feel better I'm like, "ooh!" I start rushing round trying to catch up and do the things that I'm meant to have done, you know, when I wasn't feeling so good. Oh, it is a terrible waste of time." (Julia, 1:140-54 and 348-54)

"Your body is tired, your mind's tired, you lose your concentration, you're not that great at your job because your mind it's, you know, you're just so tired / (And) even the times that you're not trying for a baby, you still want an intimate relationship with your husband, and that's completely out the window...and I just feel so sorry for him, because, you know, I've changed probably these past three months. / It has affected every part of my life." (Sarah, 1:170-84; 186-97 and 546-63)

"And the other side of it is the sex side of it...I feel so sore...you don't feel nice...(and) I've got no sex drive because I just think, "oh no, not more to do with down there."" ((Cries, blows nose)) (Clare, 1:237-48)

The heaviness of the bleed and associated pain also led many to express concerns about the possibility of serious underlying pathology and fears for their long-term health:

"The first three days are really painful. And as I say, it's also the amount of blood...You think, "oh, am I ever going to make this up again?", you know? "No wonder I feel tired, no wonder I feel weak".../ (And) every time I move I can just feel it draining away from me, I just feel it flushing away...It does, it drains you and tires you out terribly." (Ann, 1:253-64 and 88-112)

"Is the pain an indication of something worse going on down there that I don't know about? ...nobody seems to be bothered about it particularly...and this is four years now...I get quite upset about it, because you think, "well, is there going to be an end to it?" ((Cries)) / And I always feel as though I'm fighting for energy...I'm not ill, but I'm not right either." ((Cries)) (Clare, 1:222-36 and 251-6)

"It's the not knowing why it's so heavy. Why am I suffering like this?...I mean I talk about it like it's an illness, which is, you know, but it does feel like that sometimes, at certain times of the month." (Julia, 1:244-53)

5.3.2 *Symptom history*

The duration of symptoms reported amongst the interviewees ranged from three months to 33-years, with half of the interviewees having experienced heavy and painful menstruation since their menarche 20 to 30-years ago. For one of these women, symptoms had continued largely unchanged since that time, whilst the other two had experienced a deterioration in their menstrual health. Five of the interviewees had related the onset or worsening of symptoms of heavy and painful menstruation to reaching their mid-thirties, childbirth, or successive lost pregnancies:

> *"I've always been heavy-ish... (But) it's got to be the past 10 years that my periods have been very, very heavy...(And because) I get big clots that come away from me...I always think that maybes its all your childbearing things coming away from you, or whatever." (Linda, 1:31-3 and 43-8)*

> *I'd say the last five years they've been enough for me to go to the doctor and say you know, " I'm fed up with this I need something to help." / The only thing I can think of is my age... They've just progressively got worse as I've got older." (Lyn, 1:18-24 and 120-25)*

> *"About the first or second period I had after [daughter], it was, there was just so much blood, you know, it frightened me...And since then, and really since then, the bleeding has been very heavy and clotty." (Clare, 1:80-6)*

> *"You see with the first miscarriage I was fine, the ectopic I was fine, but the last miscarriage it really hit me. I think that it hit me physically as well as mentally...and I don't think I really recovered from that." (Sarah, 1:108-120)*

5.3.3 Previous treatment experiences

Three of the six interviewees had been actively seeking treatment for menorrhagia from their GP or complementary healthcare practitioner when they came to participate in ACUMEN. Two of these women had found drug therapies largely ineffective, and one had turned to acupuncture when their GP had been unable to offer an appropriate intervention, given their medical history and the risk of serious side effects:

"I went to say the Mefenamic acid just wasn't working, wasn't you know, ideal and it was actually getting worse, the pain and everything. And I went to discuss it with him, and he (GP) said, "well there is a study going on about acupuncture, would you like to try it?" (Clare, 1:354-72)

"I came out the doctors' and I just thought, "no one is going to help." I'd gone to see three different doctors over a month and no one, they just kept saying "go away" basically, all in different ways, you know, but basically "go away, we can't do anything for you, you're bleeding so what?" / That's when I started, you know, really researching other alternative methods and things, to get me back to normal." (Sarah, 1:338-356 and 356-65)

Of the three women who had not been actively seeking treatment for menorrhagia when they joined the study, two had previously experienced drug therapies as ineffective or unacceptable because of the risk of potential side-effects, and one had no experience or awareness of drug therapies to reduce excessive menstrual flow:

"As I say, these tablets (tranexamic acid) that I had from the doctor, I thought they may have worked, and as I say, they probably would have done, if I'd been able to take them properly. But it's remembering everything isn't it...(And) yes, not really wanting to take it." (Ann, 1:285-90)

"Like I say, I'd never been to the doctors about my periods really. I'd gone because I was tired and said, "I'm, I feel really, really worn down and need some iron tablets", but I've never ever gone and said, "I really have a problem with this", you know, because I honestly didn't think that there was anything that they could do."(Linda, 1:138-43)

5.3.4 Motives for joining the trial

All interviewees mentioned that they had joined ACUMEN because they wanted to find an effective and acceptable treatment for their menorrhagia. None had heard of acupuncture being used for menorrhagia, although one had heard that acupuncture might address infertility. All expressed concerns about the side effects from drug therapy, or drug therapy and surgery, and all believed that acupuncture might offer them a safe, 'natural' and low-tech alternative:

"Well it (ACUMEN poster) was in the doctor's reception...and I thought "oh, that's me", because it says "heavy periods" and "would you like help?""(Linda, 1:129-33)

"I went to see my doctor... it was my annual review...(and) she suggested I try to put my name down for this study...So I said, "yes, I'll try anything" - which is what I felt when she suggested it to me." (Lyn, 1:250-6)

The four women whose referral to ACUMEN, or use of acupuncture, had been self-initiated, expressed a strong preference for acupuncture as a 'natural' therapeutic approach:

"I like, I like the fact that it's natural. I like the fact that it's not drugs. It's not something that you take, a tablet, you know, I am averse to anything like that. I mean I don't even like taking Nurofen or anything." (Sarah, 1:671-3)

"Well, with all drugs, there's always side effects isn't there?...I keep thinking acupuncture's quite natural. You're stimulating your body to do it itself or whatever."(Linda, 1: 203-10 and 177)

"Yes, so that's why. It's not a tablet, it is a hands on thing." (Julia, 1:402)

"Again, it's not tablets. / (And) I've never heard of anybody that's had acupuncture that's made them worse...so I'll give it a try" (Ann, 1:340 and 356-68)

In contrast, the two women referred to the study by their GP, emphasised the need for alternatives to their current, ineffective management using drug treatments:

"The tablets I take (tranexamic acid) do help, but they don't help every time, and I feel as though my body's getting used to them now. When I was on the pill that was okay for a while, and then my body seemed to get used to that. So I've tried three or four different things I think." (Lyn, 1:156-60)

"The Mefenamic acid just wasn't working, wasn't you know, ideal and it was actually getting worse, the pain and everything. / I'm really desperate to get it

sorted, and if it (acupuncture) works it would be great, because it's not taking tablets." (Clare,1:354-7 and 376-80)

Positive reports about acupuncture received from friends and relatives, through books, the media and internet, had influenced all but one of the interviewees. Interviewees' confidence in acupuncture prior to their treatment appeared to stem from a number of factors including; experience of previous successful acupuncture treatment, GPs positive views about the trial, the fact that a trial was being undertaken, and to its long history, reported success and signs of growing acceptance:

"I saw this (ACUMEN GP surgery) leaflet so I thought, well it's worked for my neck so it may work for these." (Ann, 1:312)

"And I know a couple of people who've had acupuncture who firmly believe it has helped them." (Julia, 1:405-6)

"I like the fact that it's (acupuncture) an established practice, whether doctors recognise it or not, it's been around for hundreds of years. And I like the fact the success rate is great. And it's becoming more and more acknowledged by people, I mean if you've got the "The Times" writing about it." (Sarah 1:674-7)

Only one interviewee mentioned possible wider benefits of receiving treatment as a reason for agreeing to join the trial when she anticipated that a course of acupuncture would provide her with regular time "for me", away from home and her commitments as mother to four children and carer for a terminally ill husband. And for one, participation had ultimately relied upon their being a local clinic within walking distance from her home:

"I asked him (GP) how much time it was likely to take up, and he felt it was quite a chunk... I didn't know how it would fit in with the children and [husband] is away a lot ...But then I found out I could go and do it at (local street name) and it would only be an hour a week, which I could probably fit in with [daughter's] Play Group." (1:354-72)

5.3.5 Expectations of acupuncture care

With regard to their expectations of acupuncture, all of the interviewees believed treatment could bring about health benefits. Two expected treatment to address psychological and physical symptoms and promote "well-being", and two had defined the range of expected benefits according to their perceptions of acupuncture's strengths and weaknesses: Whereby they expected acupuncture to be ineffective in the treatment of symptoms arising from organic pathology, such as pain from scar tissue or infertility caused by blocked fallopian tubes, but effective for idiopathic or functional diseases, where symptoms had arisen from a hormonal or "body imbalance", perhaps as a result of emotional stress or trauma during childbirth:

> *"I've not heard of anybody that has had it that it hasn't worked for, so it's going to work, so it's going to be good." (Ann, 1:389-96)*

> *"A positive approach...but no, I don't think I can go in with high expectations, it wouldn't be fair on me." (Julia, 1:446-9)*

> *"If it's more physical, if it's more like the scar tissue or there's something not quite right physically there, then I don't think it (acupuncture) will work. But if it is to do with hormones and the balance of your body, then I think it will work." (Clare, 1:436-445)*

> *" I know there's medical things that acupuncture can't fix, you know obviously, like if your tubes are blocked no amount of acupuncture is going to unblock them...(But) I think my problem is stress related ...(and so) I think acupuncture could help me." (Sarah, 1: 915-946)*

All of the interviewees expected acupuncture to involve the insertion of needles, and two of the women had given some thought to the mechanisms of acupuncture – 'how it works':

> *"I think the needles are strategically placed. I don't know if they're placed at nerve endings... I'd presume they'd be ((holds and looks at abdomen)), but then there's so many different connections in the body isn't there? ... I mean it's like reflexology isn't it? ... the*

big toe relating to the head? I mean that's a long way from the head, but!" ((Laughs)) (Julia, 1:423-37)

Well I presume that they put the needles in at the nerve lines, nerve lines or energy lines? ... I always presume its like referred pain... If you put something in it will help somewhere else." (Clare, 1:406-415)

Based upon their own or the experience of others, half had expected to experience some discomfort, one had expected to feel dizzy and tired until her energies improved, and one had expected to be able to relax whilst the needles were in:

"[Husband] had acupuncture for a bad shoulder and he said it killed...But I know that it doesn't always hurt, it depends on where it is and how it's done! ((Laughs)) So I don't know, but I presume there will be some discomfort." (Clare, 1:388-97)

"I thought it might bleed a little bit. It did say in your leaflet actually, that you might get a little bit." (Linda, 1:240-1)

" I sat up and I nearly passed out...[acupuncturist] said, "it's all right, sometimes it does", because I hadn't much energy anyway and I hadn't relaxed at all and I wasn't used to it. / And I came home and I was just hopeless that night.... I was so drained." (Sarah, 1:565-586 and 587-592)

Three of the women had made reference to their acupuncture practitioner. Of the two interviewees yet to experience acupuncture, one had predicted that their practitioner would wear a white uniform, and one that consultations would be not dissimilar to those with their GP. The interviewee who had begun her course of treatment prior to joining ACUMEN, reported the beginnings of a strong therapeutic relationship, superior to the one she'd had with her GP:

"And s/he sat and s/he listened, which was more than anyone else had ever done. You know, actually listened...S/he was a comforting, caring person, which I liked very much... I just felt so positive when I came out of there." (Sarah, 1:494-505 and 565-586)

5.3.6 Summary

In summary, the interviews held at baseline revealed that whilst only three of the six interviewees had been actively seeking treatment from their GP or a complementary health practitioner when they came to participate in ACUMEN, they had all been experiencing symptoms of heavy menstrual blood loss and pain that had affected their quality of life and led many to express concerns about the possibility of serious underlying pathology and fears for their long-term health. Their collective experience of drug therapy had been that it was unavailable, ineffective, inappropriate given their medical history, or unacceptable because of potential side effects. A clear shared motive for joining the trial and accepting the offer of acupuncture, therefore, had been to find an acceptable and effective form of treatment. The attraction for using acupuncture had stemmed from it being perceived as a safe, natural and "hands-on", low-tech alternative to drugs and surgery. None had heard of acupuncture being used for menorrhagia. Thus, confidence in its' potential effectiveness had come primarily from a previous successful experience and positive reports of acupuncture in other contexts. Two of the women had expected acupuncture to address both psychological and physiological symptoms, and two had expected acupuncture to be ineffective for symptoms arising from organic pathology, and effective for symptoms related to idiopathic or functional disease. Half of the women had expressed some anxiety about the needles, but all had been delighted with the opportunity to receive a course of acupuncture. Table 41 provides a distilled summary of the interviewees' experiences, motives and expectations at baseline.

Table 41: Distilled summary of interviewees' reported experiences, motives and expectations at baseline

	Linda	Lyn	Sarah	Ann	Clare	Julia
Menstrual history and experiences	Heavy, painful & irregular since menarche, worse mid-30's - previous 10-yrs. Concern: long-term health Also tiredness, passing large clots, pain down legs, in fingers and wrists *(sudden flooding described at interviews 2 & 3)*	Heavy and painful since birth of 3rd child 10+yrs ago & discontinuing OCP, worse mid-30's, previous 5-yrs Concern: long-term health Also tiredness, sudden flooding, passing large clots, short cycle, mood swings, craving chocolate	Heavy & painful, then light & prolonged shortly after 2 miscarriages & an ectopic pregnancy – onset previous 3-months. Concern: long-term health, serious pathology Also tiredness and mood swings	Heavy & painful since menarche, worse for IUD no better for sterilisation. Concern: long-term health Also tiredness, pain down legs, migraine headaches	Heavy & painful since menarche, worse since birth of 2nd child 4-yrs ago. Concern: serious pathology, long-term health Also tiredness, sudden flooding, large clots, acne – Non-cyclical pain on intercourse, with bowel movements & walking	Heavy & painful since birth of 2nd child 6-yrs ago and discontinuing OCP. Concern: long-term health, serious patho Also tiredness, sudden flooding, passing large clots, pain at ovulation, bloating, poor concentration
Experience of drug therapies	None No consultations for bleed, *unaware of drug therapies - prescribed iron only*	Ineffective Regular GP consultations last 5-yrs. *3 or 4 drugs prescribed including OCP and Tranexamic acid*	Inappropriate or Unacceptable Regular GP consultations last 3-mnths - *Tranexamic acid and hormone treatments considered*	Unacceptable No recent GP consultations *OCP, HRT, Tranexamic acid prescribed*	Inappropriate or Ineffective Regular GP consultations last 4-years - *OCP, Mefenamic acid, Solpadine, Progesterone prescribed*	Unacceptable or Ineffective No recent GP consultations *OCP amongst others not stated prescribed*
Using acupuncture – motives	Surgery poster Alternative to drugs Prefers natural therapy Chiropractic & acupuncture effective for shoulder pain Acupuncture's long history Positive media reports ACUMEN evidence others have benefited	GP referral Alternative to drugs/ surgery Positive reports from friends GP positive	No conventional treatment Alternative to drugs & IVF Prefers natural therapy Acupuncture's long history Positive reports (books, media, internet, niece) Prefers holistic to reductionist approach	Surgery poster Alternative to drugs /surgery Prefers natural therapy GP acupuncutre effective for neck pain Positive reports from friends Sessions time-for-self	GP referral Alternative to drugs Rebalance / restore normal body function Local clinic	Surgery poster Alternative to drugs Prefers natural therapy Positive reports from friends
Using acupuncture – expectations Effective? Treatment process	Could be effective Needle insertion at specific sites. Needle insertion may cause bleeding	Improved well-being Can improve psychological and physical symptoms Needle insertion at specific sites, practitioner wears a white uniform	Symptoms from organic pathology = Ineffective Symptoms from idiopathic/ functional diseases = Effective (e.g. mind/body disharmony, menstrual & digestive problems, stress & low energy post emotional stress for lost pregnancies) Dizziness/ tiredness after needles if energy low Superior therapeutic relationship to that with GP	Will be effective Needles inserted locally Retained for 20-mins and stimulated every 10-mins Possible discomfort	Symptoms from organic pathology = Ineffective Symptoms from idiopathic/ functional diseases = Effective (e.g. hormone/body imbalance post trauma in childbirth) Needles at specific sites along nerve / energy lines Distal points of benefit – like referred pain. Discomfort, practitioner like GP	Could be effective Needle insertion at specific sites - related to nerve endings? Points distal to uterus can have an effect on uterus – like reflexology. Relax with needles in for some time

5.4 Exploring the meaning of treatment acceptability

The following is a thematically organised account of the meaning of 'treatment acceptability' to the women participating in interviews and receiving acupuncture for menorrhagia in primary care for the purposes of ACUMEN. This account does not always distinguish between the data from interviews held at three month and six month follow-up, as analysis revealed an almost seamless continuation of the patients' story, where views held at three months about the appropriateness of GP referral and the acceptability of acupuncture were not disparate from those expressed at 6-months.

First, each core theme is described and illustrated. Then an interpretation of the follow-up data set as a whole is given, with reference to the findings from the descriptive analysis of the initial baseline interview: The patterns and connections between ideas, views and experiences are mapped-out to illustrate the meaning of treatment acceptability to the interviewees. This thematically organised account is given under the following headings *[see also Table 38]*:

I. A safe, natural alternative to drugs and surgery *[5.4.1]*,
II. The acupuncture treatment process *[5.4.2]*,
III. A holistic perspective *[5.4.3]*,
IV. Perceived outcomes from acupuncture treatment *[5.4.4]*,
V. Patterns of use – conventional medicine and acupuncture *[5.4.5]*, and
VI. Mapping and interpretation: an exploration of the meaning of treatment acceptability *[5.4.6]*.

5.4.1 A safe, natural alternative to drugs and surgery

By the mid-point of their course of ACUMEN acupuncture treatment, after approximately 10 weeks of once weekly sessions, all the interviewees had commented on how they had enjoyed having access to what they perceived to be a safe, natural therapy and saw this as a key aspect of treatment acceptability *[theme chart one, Appendix 45]*:

"If you go to the doctor...they're going to give you a pill that stops the heavy bleeding, and surely, isn't it better to go somewhere for acupuncture that does it naturally, so you're not putting a pill into your body anyway?" (Sarah, 2:626-35)

"I think (women) they'll like it because they'd feel it is maybes a way of treating their problem without having to go to surgery, which is what I like it for." (Lyn, 2:277-8)

"In a way I've chosen to do the acupuncture. Well, you've given me the opportunity, but I've chosen to do this as well. You don't get a choice like that at the GP, it's like you either take this tablet or you have this treatment, that's it...it's about feeling in control, yes." (Julia, 2:128-40)

5.4.2 The acupuncture treatment process

The actual experience of attending for treatment, the sensation from the needles, the environment and nature of the therapeutic relationship, along with the time spent getting to and receiving treatment, were all factors that determined the interviewees' willingness to attend for acupuncture *[theme chart two, Appendix 46]*.

5.4.2 a. "Just a little pinprick"

Most of the interviewees believed that a fear of needles is common, and that concerns about discomfort and pain are, therefore, a key determinant of treatment acceptability:

"They (people) always ask me if it's painful, and things like that." (Lyn, 2:256)

"They shouldn't call them needles should they, because they're so fine, they shouldn't put people off." (Clare, 3:497-9)

Following their experience of acupuncture, these women had wished to stress how painless needle insertion had been for them, especially given that any discomfort was momentary and preceded a pleasant period of relaxation and a profound sense of calm and well-being that often lasted for some time following the treatment. The discomfort upon needle insertion or when stimulating the needle to give the 'deqi' sensation was also considered to be 'worth it' in light of health benefits:

"There's just a little pinprick and then that's it. It doesn't hurt at all." (Sarah, 2:88-96)

"...It's seconds really... they stop hurting as soon they're in. / ... I thought, "oh, they're sticking needles in you, it must hurt." It doesn't, and you have this feeling of well-being when you come out." (Lyn, 2: 72-84 and 3:194-6)

"I think there's lots of people that are nervous about needles and things...it's making that person realise, convincing them, that the benefits far outweigh the slight discomfort that you might have while having the treatment...I mean it is a bit unpleasant having the needles, and I did, I was sometimes reluctant because sometimes they can catch and hit something, and it's like, "oh my goodness!" / ...But gosh there's a price to pay for everything isn't there? And it's a very small price to pay if it's working." (Julia, 3:181-95 and 2:522-3)

5.4.2b "You feel so relaxed"

All of the interviewees made particular reference to how much they had liked and valued the relaxation that came from receiving acupuncture. For one of the women, this seemed to be related to environmental factors - the people, smells and sounds within the clinic - whilst the others associated the relaxation and calm they experienced to the acupuncture itself:

"I've appreciated going to a pleasant, relaxed environment where people are very, extremely pleasant. / You'd go in and smell fresh coffee, and sit on nice chairs, and read nice magazines, and yes, a nice carpet and pictures on the walls, and really good treatment rooms. / I mean it's nice to go in and lie down and listen to classical music." (Julia, 3:138-9 and 3:153-62 and 3:152-3)

"...Afterwards I'm so relaxed, completely and utterly relaxed / (And that's) so important when you have a heavy lifestyle. / Acupuncture does give you this wonderful feeling of relaxation." (Sarah, 2:101-6 and 2:232-4 and 2:272-6)

"...When you're having the treatment you feel as if you're drifting, you feel so relaxed. / ...You go in with all the world on your shoulders, and I know s/he can't get rid of all your problems, but they seem so minor when you come out, because you do feel so relaxed...you take it in your stride." (Ann, 2: 31-47 and 2:185-97)

"What I wasn't expecting was the feeling of calmness, you know, while you're having it, that nice warmth and just sort of out of it feeling - really nice." (Clare, 2:134-6)

5.4.2c "Someone listening, someone understanding"

All the interviewees had valued the relationship they had with their acupuncture practitioner, and this had stemmed from feeling listened to and understood at both the initial and subsequent consultations. This often had the effect of validating their experience of ill-health, and many of the women had described how it had engendered their trust in their practitioner's ability to prescribe appropriate treatment:

"I think what helped was someone understanding, someone listening for that amount of time. You get 5 minutes with a doctor if you're lucky." (Sarah, 2:262-8)

"S/he got a true picture I think...s/he knew before s/he started what I was like really. Yes, s/he's really good." (Lyn, 2:144-7)

"They (women) would like it because somebody shows an interest and they listen to what you've got to say. / When you go to see the doctor it's just a rush. The acupuncture is, "how has it changed? And lets work on how it's changed", you know, they take notice of what has happened...delve into it a bit more, ask questions around it to try and get the full story." (Clare, 2:476-7 and 2: 232-51)

Their practitioner's ability to accurately interpret physical signs of health, or seemingly to intuitively locate areas of disease, often provided further assurance of their healing skills:

"...she knew where I was having trouble in my joints and things like that...so she knew a lot of things that I hadn't told her." (Linda, 2:133-6)

And all of the interviewees' had appreciated their practitioner's calm demeanour, friendly nature, or had derived reassurance and hope from their practitioner's confident approach and willingness to help:

"If you've got a nice acupuncturist it does make a difference, because as I say, mine is cool, calm and collected. Very calm, you know, don't look as if s/he's got anything on her mind, and everybody must have something. Everybody! ((Laughs))" (Ann, 2:476-9)

"...S/he had such confidence in her/his medicine that that gave me confidence, because s/he never once wavered, you know." (Sarah, 2:377-85)

One of the interviewees had described how she had valued feeling cared for, and two had referred to the time spent with the practitioner over a period of some six -months, and how this had led to the development of a trusting, supportive, therapeutic relationship, where the patient's health becomes a shared responsibility. Two of the interviewee's had also noted how considered they had felt when arranging appointments:

"You're treated so well, and it's like someone really cares about my health, you know, apart from, someone apart from me and my family, somebody else cares about it. And that makes you feel better again." (Julia, 2:116-9)

"Also the fact that they are more sympathetic and listen to you a lot more and help with the problem on a regular basis. A doctor will listen and say, "okay we'll give you this tablet", and then that's where it ends, isn't it?" (Clare, 2:639-42)

5.4.2d "It does take time"

The time spent getting to, and receiving some 20 treatments over a period of six months or so, led the interviewees to discuss the opportunity cost involved for themselves and the issues they imagined other women might face when deciding whether or not to embark upon a course of acupuncture for menorrhagia. On the one hand, the women recognised how they had benefited from regular time-for-self, and how this might also appeal to other women:

"...And it's time-out as well, it's great! You know, that is allowing me to go and, it's saying, "yes it's okay Julia, for you to go to this place, lie down, and relax for half an hour", which is like, "yes, okay, I don't mind that!" ((Laughs))" (Julia, 2:347-51)

"Women might be put off because it does take time to go and do it, but having said that, it's not time that is horrible, it's just that you've got to take time off in the day to do it. So it's alright for me that's not working...but then it might be quite nice after you've done a days work to come and have an hour of quiet and calm." (Clare, 2: 481-500)

However, they were also very aware that this was time that could be spent elsewhere, and that many women might find attending for treatment very difficult, given their responsibilities:

"I can't say there's anything I dislike about it, except for the time it takes... that's all right at first, but then after a while you start thinking, "oh, you could be doing this, you could be doing that"... because it's a long drawn out thing, it's not an instant fix." (Linda, 2:376-86)

"Just like it actually takes up time. You can't just take a tablet and get on with your day. You've got to take time, make time to do it. / And if you've got other kids to look after, well, what happens to them then?" (Clare, 2:478-9 and 2:481-500)

Two of the interviewees described how they considered that time well spent, given the benefits they had experience, but half believed that where benefits were less immediate, adherence to the full course of acupuncture could prove a considerable problem in any future studies, and they wished therefore, to stress the need for patience:

"It was a slow process. So give it time...I mean it's hard to keep going, you know, through something when it isn't, you know, "oh it's not working I'll give up", but I just say, "give it a chance to work." (Lyn, 3:213-25)

5.4.3 A holistic perspective

The holistic medical paradigm that underlies acupuncture proved to be a key aspect of treatment acceptability. Interviewees spoke at length about how they had liked and valued the fact treatment had taken into consideration the full array of physiological and psychological symptoms they experienced, rather than focusing upon an isolated

symptom, and that it had involved giving practical advice about ways they could help to bring about and maintain health *[theme chart three, Appendix 47]*.

5.4.3a "Feeling better completely, as a whole"

For five of the six interviewees, the fact that the consultation and treatment had not simply focused upon their menstrual bleed was an unexpected but important aspect of treatment acceptability. They had liked that their points prescription and the use of adjunctive techniques, such as moxabustion and acupressure-massage, might change from week to week to attend to distressing symptoms such as migraine, neck and back pain. And in one case this had the affect of raising health expectations:

> *"...s/he was interested in all aspects, with all of me. So, so, not just "how've your periods been"? ((Laughs)) (Ann 3:167-70)*

> *"I think it's just that acupuncture covers more than a pill does. That is the big thing for me."(Sarah, 2:642-5)*

> *"...it's made me realise that maybes you don't have to put up with these things, you know?"(Linda, 2:79-82 and 2:196-207)*

Five of the women had also described how they had enjoyed learning to understand themselves and their symptoms from acupuncture's holistic perspective: the meaning and interconnectedness of different signs and symptoms, and the impact of lifestyle and emotions upon physical function.

> *"But the best thing, and I read it in the newspaper article, the best thing was that it was somebody talking to you about the whole, your whole system, not just one symptom...And it was good because it made me think about it as a whole as well. / So (acupuncturist) s/he's changed my view on that, in that I've got to be healthier, generally healthier first, before the periods can start to get better." (Clare, 2:30-9 and 2:187-95)*

> *"...it's talking and it's finding out about yourself. It's a whole different experience than what you actually expect, I think.... It's not just going there and getting jabbed, you know?" (Linda, 3: 253-77)*

"...I know stress could affect the period, and that if something happened you can miss a period. I knew that, but not that a hectic lifestyle would, or eating wrongly, or eating the way I was, or anything like that. No I didn't, I just didn't..." (Sarah, 2: 302-15)

And more than half of the interviewees felt the fact that benefits from treatment could be far reaching, to be an important aspect of treatment acceptability that should be communicated to women considering a course of acupuncture for menorrhagia. They also felt that knowing that acupuncture would be of some benefit, even if not to the menstrual bleed, would positively affect women's willingness to receive acupuncture:

"...give acupuncture a go because not only does it help the problem, the period, the heavy bleeding, but it will help you as a person. It will give you relaxation, it'll make you stress free, it will give you a good feeling about yourself. / ...acupuncture does get to the root of the cause and it does make you feel better completely, as a whole." (Sarah, 2: 633-7 and 3:325-40)

"...it's not just the period that it will sort out, and that's the benefit of it, it treats your whole system, it's all the other added benefits, so okay you might say, "well I can put up with this period pain or this discomfort or whatever, so I won't bother to make the effort to go to acupuncture", but it is worth making the effort because it's all the added benefits that you get, it keeps you healthier... it will improve your quality of life, it won't just improve your periods, whether it does or it doesn't work for your periods." (Clare, 3:468-87)

5.4.3b "The diet thing as well, that's been really good"

Five of the six interviewees had been given practical advice about ways they could help to bring about and maintain health. This advice had ranged from being encouraged to recognise and assert their own needs, to making simple changes to their diet - such as eating cooked foods, or reducing their coffee intake.

"S/he put it to me in that its your body has to heat up the food to break it down when it goes in, so if it can go in hot your body doesn't have to put that energy into it to start breaking it down...and you get more goodness out of the food... / ...s/he put me on porridge for breakfast and soup for lunch, and that seems to have done me a lot of good,

definitely, in that my energy has gone up and I keep going better." (Clare, 2:152-71 and 2:173-80)

"S/he's very good on advice...s/he's telling me exercises and things... / ...cutting down on my wheat intake.../ (And) s/he's trying to make me more assertive I think.../ ...I'm taking it all on board, and I'm trying to lead a little bit different lifestyle wise." (Linda, 2:91-110 and 2:261-70 and 2:273-6 and 2:276-96)

All the women had found the advice insightful and of value, and often commented that this was an aspect of treatment that they had not expected. And for one of the interviewees the insights given into the relationship between behaviour, emotions and health empowered her to make quite profound changes with regards her perspective on and approach to life:

"The diet thing as well, that's been really good. It's these benefits I didn't expect from acupuncture. I thought you'd just go and have the acupuncture...a bit like going to a chiropractor; you'd just get that bit sorted. But it was the diet, go and try this diet to improve your whole, your whole body health." (Clare, 2:526-32)

"And all this from a little bit of acupuncture eh?" (Linda, 2:298)

"...I just thought that it was good that I didn't listen (to my body), and that I just did these jobs like superwoman...And so it's made me realise that, you know, you've got to have it in proportion.../ (So) I was helping with the food, and I was helping by not doing too much at work, and certainly the stress dropped straight away. The first week the stress stopped, because I made a conscious effort that I wasn't going to be that stressed anymore. / It's not just about her/him treating me, it's about me helping my own lifestyle as well." (Sarah, 2: 281-97 and 2:120-36 and 2:357-60 and 2:360-2)

5.4.4 Perceived outcomes from acupuncture treatment

The outcomes from acupuncture included changes to the menstrual flow; changes to other symptoms related to the menstrual cycle, in particular pain and energy; changes to symptoms not specifically related to menstruation, such as digestion and shoulder

pain; changes to their general health and well-being; and changes in health behaviour *[theme chart 4, Appendix 46]*.

5.4.4a "It's manageable now, you know, I can cope"

All the interviewees understood that despite the holistic focus, the primary objective of acupuncture treatment was to normalise menstrual blood loss, and at three months follow-up, two of the women had described their bleed as "normal", and four had described how their bleed, whilst still heavy, had improved by way of a shorter bleed, less flooding, or passing fewer large clots.

> *"Well it's just incredible...a natural period I suppose... in fact, to me I thought, "hang on, we've gone too much the other way here, I've dried up!" ((Laughs)) / ...almost don't realise you're having a period because it's no big deal." (Julia, 2:183-223 and 2:228-47)*

> *"...my periods are still heavy, but I've noticed this time I didn't have to get up during the night. Now normally I'd be up three, four times during the night to change... /...(and) it's not just your pad you change at night, you know, it's like everything...I mean I still changed, I went through four pairs of trousers in one day...(but) I felt better for having a full nights sleep..." (Linda, 2:33-48 and 2:356-69)*

> *"So the fact that the cycle's lengthened is a big plus, and the flow has got shorter too... /...(and) I haven't had big clots." (Lyn, 2:334-8 and 2:51-64)*

At six months follow-up and the end of the treatment period, benefits had been maintained for the two interviewees whose menstrual loss was normal at three month follow-up, and two women reported that they were satisfied with the outcome achieved; for whilst they still experienced heavy menstrual loss, it was for fewer days, there was no sudden flooding, and they no longer used prescribed medication to control the bleed. However, two of the interviewees reported that they had had very little or no respite from their symptoms of heavy blood loss over the six months of treatment.

> *"...I mean mine are just so easy now, and so light." (Sarah, 3:108-17)*

"I would say it's heavy for three days and it's like continuous blood for three days, very bright, but it's manageable, because there's no big clots, and no stopping and starting... it's not flooding.... It's something like five or six days now, or three days heavy and a couple of days mediocre, and that's it, whereas before it was stretching to a week or more." (Lyn, 3:45-62)

"But yes, it's worked for me. / I had two really heavy days and then the other four were light and was only on six days..." (Ann, 2:52 and 3:78-91)

"The periods haven't got any better in terms of flow. They've been heavy, and heavy enough to worry me...with the flooding and clots and fibrous material. ((Cries))" (Clare, 3:47)

"I've just finished a period. It was very heavy. So heavy I was having to wear incontinence pads. So heavy I was like trousers, everything ((indicates flooding))...So heavy I marked my friends settee. ((Upset and angry))" (Linda, 3:7-26)

5.4.4b Other symptoms related to menstruation

By the end of the treatment period, all six of the interviewees reported that they experienced less pain at menstruation, five reported improved energy, three described how their cycle had also become more regular, and three reported feeling calmer, or less "ratty", both prior to menstruation and in general.

For all of the women a reduction in pain had been a valued outcome that had improved quality of life, and for one, finding effective pain relief had been the priority at the outset of treatment:

"And yes, I do get a bit of discomfort, but nothing compared to what I used to have. I mean, my God, I used to go to bed for the afternoon.../...a bit of tummy ache now, but I think, that to me, that's natural." (Julia, 2:259-61 and 3:81-104)

"...fighting the pain, I mean that was the worst part.../ (So) the priority for me in the beginning was the pain, and the pain's gone. I don't take much in the way of painkillers at all now, whereas six months ago I was writhing on the bed in agony, and that's gone." (Clare, 3:72-76 and 3:449-452)

Extreme tiredness was a symptom that all the women had shared when they joined ACUMEN. At follow-up, five spoke specifically about how treatment had improved their energy levels and for four of the women, this change had been such that it had had a considerable impact upon their quality of life:

> *"...before this [study]...I felt pretty depressed and tired, and I felt that I wasn't doing anything, and it was like, "oh God, you're not asleep again?"...People just sort of say things, and I thought well I can't...I just haven't got the energy. But I have now, I have...I feel like I'm on an even keel now. Oh, I'll start to weep in a minute! ((Laughs and cries))" (Julia, 2:393-427)*

> *"...although I get tired, I don't get desperately tired...maybe a couple of days a month I might have to go and lie down, but I'm not like every day having to go up and lie on the bed just to get to the end of the day. So that's a big plus." (Clare, 3:426-443)*

> *"I have longer days in between where I'm full of energy, whereas before... it was like, yes, a relentless cycle." (Lyn, 3:81-9)*

The three patients who reported experiencing more regular menstrual cycles in response to treatment enjoyed having more time between cycles, feeling physically ready for conception, or being better able to plan events, such as holidays, not to coincide with menstruation.

Three of the interviewees described feeling less "ratty" prior to menstruation and far calmer generally. One of these women commented that her family had appreciated her shouting less. And for one of these women, this was a most important treatment outcome, as their irritable, erratic and uncontrollable moods were having a negative impact both upon their sense of self and their relationship with their husband:

> *"After about four or five weeks I noticed a complete change, in that I was more level, more calm.../...s/he must have balanced my hormones as well...and for me it was the mental really that was affecting me as a person more than anything else. You know, because if you feel you've lost control of your life that's horrible./ ...when I look back I just feel so sorry for him really (husband), because he was just living with this monster..." (Sarah, 2:385-399 and 2:414-34 and 3:175-194)*

5.4.4c Symptoms unrelated to menstruation

Of the four interviewees who had found the acupuncture they'd received in ACUMEN to be of benefit for symptoms unrelated to menstruation, three had experienced improvements to musculoskeletal conditions - neck and back pain, shoulder or hip pain; two had described being less stressed; one had reported improved digestion and immunity (catching colds); and one improved bowel function. And it is worth noting that included here, are the two women whom had not responded to treatment to reduce their menstrual flow within the six month treatment period:

> *"...I had a shoulder complaint...and it worked a treat...s/he just knew where to get to the point that s/he needed to get to – and it was fabulous. I've had, in fact I've had no trouble with it since..." (Linda, 3:321-41)*

5.4.4d "The treatment I had had an affect on all of me"

All of the interviewees reported that acupuncture had had a positive affect upon their general health and well-being, and one had described how such improvements had better enabled her to cope with her excessive menstrual loss:

> *"I think, I think acupuncture has done a lot for my general health...And it got rid of the pain very quickly... And it's done wonders for my hip as well, so I think for my general health and the pain it's been great. / I'm not as, it's not such a black hole, it's not as worrying, I feel as though something's been done...(And) because my general health is better I can cope with the, the thought of it, the thought of the horrible bleeding every month and although I get tired I don't get desperately tired... So that's a big plus." (Clare, 3:255-64 and 3:426-443)*

Two of the women had also referred to the positive effect simply participating in ACUMEN had had upon them, as it had validated their experience of poor menstrual health and given hope by way of a new therapeutic intervention:

"...being involved in this was just absolutely great...that makes you feel so much better...the fact that somebody's actually said, "right, we're going to try and do something about this," makes a hell of a difference." (Julia, 2:108-114)

However, it was the two women who had experienced a normal menstrual flow in response to acupuncture, that had experienced the most profound shift in their experience of health and, as a consequence, their perception of self:

"... it just seems a bit bizarre in a way. I've had these needles put in and become a changed woman! Can you imagine if you said that to somebody? "Oh right, yes!" ((Raises eyebrows, laughs))...(But) it was like two weeks out of every month I felt rotten, in one way or another, whether it be feeling tired, or feeling ratty, or being bloated and having pain, and then bleeding and having the big clotting, and everything that went with that. And it was just like, it got you down, and then you had sort of like, a couple of weeks respite to recover...and then you'd be back on it again! ((Laughs)) ...It's because I don't, I'm not experiencing all those terrible things that were related to my menstrual cycle, that I'm well, normal." (Julia, 2:275-289 and 3:81-104)

5.4.4e "I'm making a conscious effort to keep healthy really"

Changes in health behaviour was not a predicted outcome, and yet one of the women reported that treatment had caused her to think more about ways to maintain her health, and she had therefore invested in food supplements; and two of the five interviewees who had received lifestyle advice from their acupuncture practitioner, described how this was continuing to influence their behaviour, now that their course of acupuncture had come to an end:

"...it wasn't just an illness...it affected all of my life...and I never want to be like that again.../ (So) I do watch now, you know, I do make sure that I do eat properly...and that there's no stress in my life, and that if there is, then I take a different direction, which is all to do with the advice I got from this [study]...now I am healthy I want to keep healthy, you know, for me." (Sarah, 3:202-213 and 3:143-157)

5.4.5 Acupuncture and conventional medicine – patterns of use

The way in which interviewees had used acupuncture and conventional medicine over the course of the trial, what they wished they'd done in hindsight, whether they'd recommend acupuncture to other women, and how they would wish to use acupuncture and conventional medicine for menorrhagia in the future, provided useful insights into the interviewees' perception of treatment acceptability and the role they would consider for acupuncture in their healthcare *[theme chart 5, Appendix 47]*.

5.4.5a Acupuncture as adjunct or alternative to conventional medicine for menorrhagia

On joining ACUMEN, two of the six interviewees were using prescribed medication for their menstrual loss. One discontinued their prescription of tranexamic acid within the first three months of acupuncture treatment, as advised by their acupuncture practitioner, and the other continued to consult their GP and use prescribed medication alongside acupuncture. Thus, acupuncture was used as an alternative to conventional medicine by five, and as an adjunct to conventional medicine by one of the women participating in interviews. At the six month follow-up interview, all the interviewees stated that they were glad that they had accepted the offer of acupuncture as part of ACUMEN:

> *"I've found it great. "Fantastic" in a word! ((Big smile and laugh)) Oh yes, it's been a good move." (Ann, 2:382-4)*

> *"(If I could go back in time six months, and knowing what I know now about acupuncture) I would have it, you know, if it was going to make me regular and if the achy pains do stay away in my legs, that's a good thing, but I'm disappointed my periods are still heavy." (Linda, 3:31-36)*

Only one interviewee would have done things differently if they'd had the benefit of hindsight, and that would have been to use the MIRENA coil alongside acupuncture, as opposed to the range of drug therapies they had tried over the course of the study - the MIRENA Coil had been prescribed by a Consultant Gynaecologist for suspected endometriosis just prior to six month follow-up:

"I think acupuncture has done a lot for my general health, and I don't think the MIRENA coil would have done that. And it got rid of the pain very quickly...And it's done wonders for my hip as well, so I think for my general health and the pain it's been great. But I have to say that I'd have the MIRENA for the bleeding because it's so heavy, but whether it works." (Clare, 3:255-65)

5.4.5b Predicted future use of acupuncture and conventional medicine for menorrhagia

All four of the women who had found acupuncture effective for menorrhagia asserted that they would wish to return for another course of acupuncture, should their menstrual symptoms return:

"If it came back as bad as it was before, I would think right, I'll go and try another course of treatment first, because it did address all the problems, so it obviously works for me." (Lyn, 3:145-147)

Whereas the two women who had not experienced relief from symptoms of heavy menstrual blood loss, said that they would not agree to another course of acupuncture, and that their next course of action would be a new drug therapy, as both continued to wish to avoid surgery:

"...I don't think I'd have acupuncture again... being as heavy as I am, I think, you know, I wouldn't go through, I wouldn't, even though it was an enjoyable experience, I don't think, I don't think I'd go and have it done again if I was given the opportunity." (Linda, 3:44-53)

"...I don't think acupuncture is going to solve everything.../...(so) they've recommended that I have the MIRENA coil.../...(and) I suppose the ultimate is that I don't want to have a hysterectomy..."(Clare, 3:101-105 and 3:107-120 and 3:386-422)

And it is notable that whilst one of these women had been using drug therapies since menarche, the other had been first, unaware prescribed medication was unavailable, and second, attracted to acupuncture as she disliked drug therapies – *"I wouldn't feel like I was poisoning my body" (Linda, 1:203-10).* However, by the six month follow-

up she stated that she would be willing to 'try anything' to gain relief from excessive menstrual loss:

> *"Well, if there was, if there was anything on the market that would make it lighter, yes, I think I'd try anything. / (But) I don't know about surgical...that would be the very last resort." (Linda, 3:60-4 and 3:77-83)*

5.4.5c "How long is it going to be before it all wears off?"

The single concern expressed by all the interviewees about acupuncture was whether the health benefits achieved during ACUMEN would be lasting, or whether their health would deteriorate without ongoing treatment:

> *"...I don't want it to wear off, otherwise I'm back to square one." (Ann, 2:526-531)*

> *"I'm slightly bothered my energy levels will go, will disappear again, once I stop coming for the acupuncture..." (Clare, 3:312-327)*

The four interviewees who had found acupuncture effective for menorrhagia, and favoured this intervention over conventional medicine, had discussed the likelihood of long-term benefits with their ACUMEN practitioner prior to the end of their course of treatment and six month follow-up, and agreed the role acupuncture might continue to play in their health care. Three of the four were surprised but pleased to be informed that where benefits were not maintained, they would likely need only an occasional 'top-up' – perhaps twice yearly, or at times of stress:

> *"Just maybe a top-up now and again. But no, as far as [acupuncturist] is concerned, [acupuncturist] was saying that's it, you know, basically you're sorted. So the treatment's been done, and that's it, which I just find absolutely incredible, really, astonishing. /...I mean, if I have to go and see [acupuncturist] once every few months, that doesn't matter, that's fine." (Julia, 3:32-37 and 3:50-54)*

> *"...s/he [acupuncturist] said I might just have to have maybe one or two a year, just to keep things level..." (Lyn, 3:122-125)*

And one woman had determined through experience, that to maintain the full array of health benefits achieved, she would need monthly acupuncture and was delighted that this would continue to be possible outside of the trial:

> *"And I'm staying on! ((Laughs)) Oh yes, can't do without it...and it's nice to know that [acupuncturist] is there. / I went 6-weeks without any acupuncture...(and) I could feel myself uptight inside...(and) I'd shouted...So I was ready for going back...I can't go 6-weeks without..." (Ann, 3:14-19 and 3:127-153)*

Indeed, all four of these women described how concerns about their future health had lessened now that they had found an effective therapy for menorrhagia that they could access at-will, and often described how they valued knowing that they would continue to have their acupuncture practitioners support outside of the trial, should they need it:

> *"I'm always frightened my periods will go funny again.../... but I've always got that, you know, that if it did happen, I would just go back to [acupuncturist], and I know that it would be sorted. I mean, within a month I think I'd be back on track, you know, because that's what happened before...And it's nice that there's always someone, there's always help out there." (Sarah, 3:221-229 and 3:233-44)*

5.4.5d Predicted future use of acupuncture – willingness-to-pay

At six month follow-up, five of the interviewees talked about the role they perceived for acupuncture in their future health-care. The type of role envisioned stemmed directly from their experience of acupuncture in ACUMEN and the range of health benefits received. Thus, the two women who had found acupuncture to be effective for menorrhagia and symptoms such as poor digestion, migraine, back pain and stress, perceived a much wider role for acupuncture in their future health care than either the two women who had only experienced benefits to menstruation, or the woman who had not gained relief from heavy menstrual loss, but had experienced improvements to other symptoms such as shoulder pain and energy levels. Indeed, one of the women experiencing a wide range of benefits asserted that their acupuncturist would now be their primary carer, and they would use their GP only in instances where acupuncture had nothing to offer:

"I'm not giving up on doctors. I think they're very good for certain things.../ Yes, what acupuncture couldn't cover, I would then go to the doctor's." (Sarah, 2:566-584 and 2:494-505)

However, this predicted future use of acupuncture outside of the trial has cost implications, which therefore meant that these five interviewees came to express their perception of treatment acceptability in terms of their willingness-to-pay out-of-pocket for future acupuncture:

"...what puts me off, and a lot of people like me, is the cost of it... I mean at the moment I'm loving it, because I'm thinking I don't have that worry, you know? ...(But) to actually continue with that if say, you know, once these trials have been finished and that...that's an awful lot of money really, to keep forking out every week..." (Linda, 2:141-158)

"...if I'd had to pay for it I would have done, I would have paid for it, but it would have been, certainly for me, it would have been a big chunk out of my money, you know. I'd probably have had acupuncture and not have been able to buy clothes for six months, you know, because it is a lot. People pay £20, £30 for a session, and they have one session a week, you know, that would be the only draw back for normal working women like me." (Sarah, 2:681-6)

And it became clear that the interviewees were willing-to-pay for treatment where they could be relatively certain that the desired health outcomes would be achieved:

"I probably can't afford it on a weekly basis, and I'd rather not ask how much it is, but if it would carry on helping the periods by going once a month, then I'm sure that would be ideal." (Ann, 2:499-511)

"...I mean I had a shoulder complaint ...and it worked a treat. ..in fact I've had no trouble with it since... (So) I think it's great for injury and that type of thing. Yes, I would have it for that, even with the cost. / ...because I've had that really positive thing from [acupuncturist] I would definitely, definitely, if I had that complaint again I wouldn't go to a chiropractor, I'd go to an acupuncturist. Yes." (Linda, 3:321-341 and 3:367-71)

5.4.5e "Oh, go for it!"

Based upon their experiences of acupuncture in ACUMEN all of the interviewees asserted that they would recommend acupuncture to other women with similar menstrual symptoms. Indeed, one had a friend who had joined ACUMEN on their recommendation; one had instigated their sister and niece receiving acupuncture for menopausal symptoms and menstrual irregularities respectively; and two had mentioned the benefits they had received from treatment to friends:

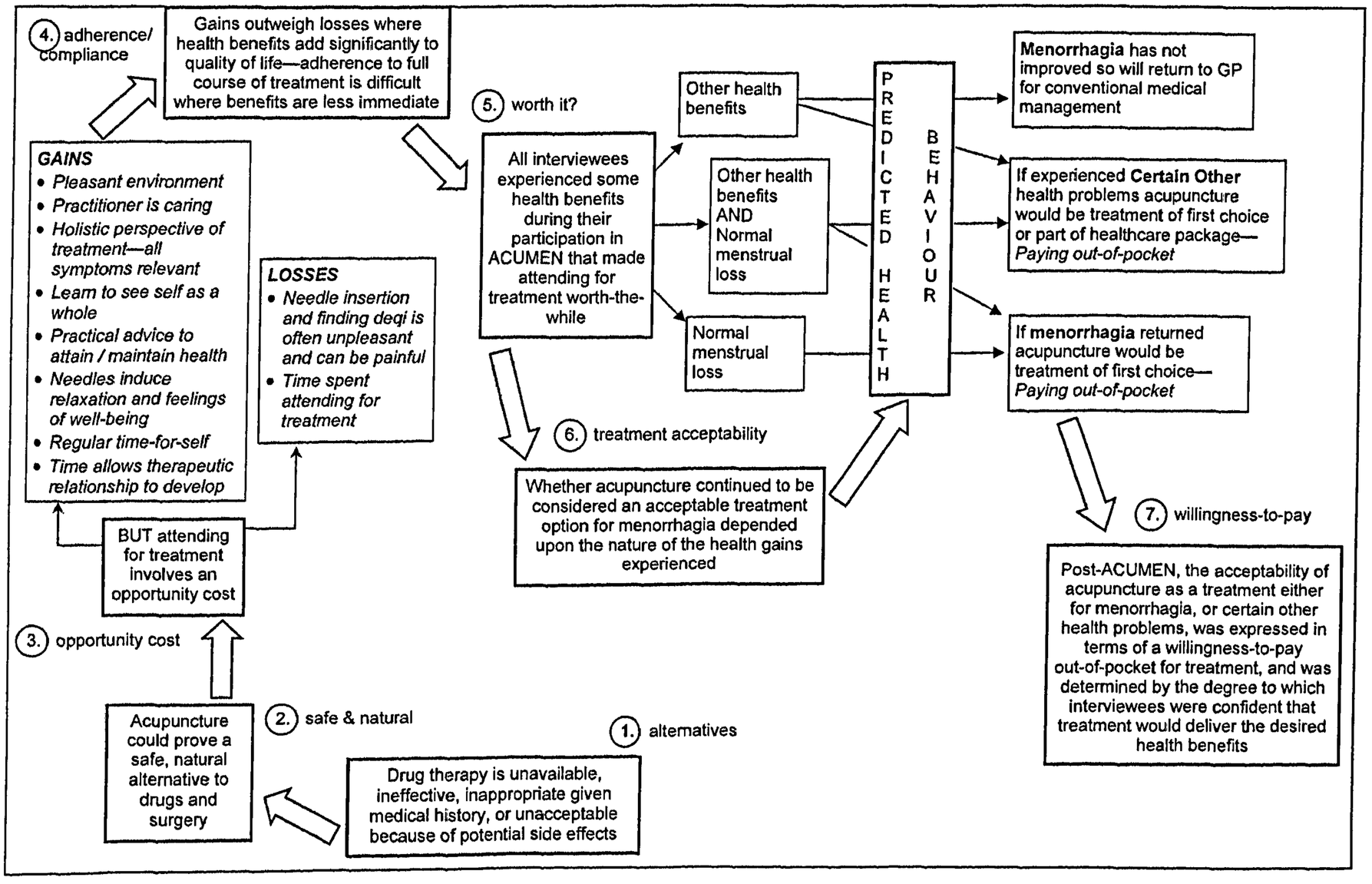
> *"Well if I had a friend who was suffering with bad periods, whether it was heavy or whether it was painful or irregular, I would always, always tell them to go for acupuncture now." (Sarah, 2:622-4)*

However, half of the interviewees noted that the people they'd spoken to had not always seemed very interested, and they were therefore unsure whether acupuncture would always appear a credible treatment. They therefore advised that future trials consider how to overcome the fact that acupuncture's potential role within healthcare is little known:

> *"...I mean, obviously I've mentioned to friends that I've been doing this, having acupuncture, but nobody's, you know, I don't think people know an awful lot about it to be able to comment to be honest. I think that's what it is, because they've not really said very much at all. So yes, people don't know. (Julia, 3:175-9)*

> *"...I do know other people where you'd meet the response, "what's a few needles in you, what does that do, how can that make you feel better?" And they don't understand, you know, and that's going to be quite difficult, you know, to get over to those kinds of people.../...(because) there hasn't been that much information on it (acupuncture), it's only been coming into its' own in the last couple of years really..." (Sarah, 3:356-62)*

Fig 8: Mapping and interpretation: exploring the meaning of treatment acceptability



5.4.6 *Mapping and Interpretation: an exploration of the meaning of treatment acceptability*

The above theme-by-theme account is useful in defining and illustrating those factors found to be of relevance to 'treatment acceptability' for the six ACUMEN interviewees. However, if we are to better understand the meaning of 'treatment acceptability' for this group of women, it is necessary to interpret the follow-up data set as a whole, and with reference to the findings from the descriptive analysis of the initial baseline interview, in order to discern the patterns and connections between ideas, views and experiences that all comprised and gave rise to the interviewees' lived experience of 'treatment acceptability' over time. The findings from this analytic process of mapping and interpretation are illustrated in Figure 8 and described below, as a series of seven cognitive steps that chart their journey from joining the study to the end of the treatment period:

Step One: alternatives

On joining ACUMEN, all the interviewees' were experiencing symptoms of blood loss and pain that had affected their quality of life and led many to express concerns about the possibility of serious underlying pathology and fears for their long-term health. Their collective experience of drug therapy had been that it was unavailable, ineffective, inappropriate given their medical history, or unacceptable because of potential side effects. Thus a clear, shared motive for participating in the study and accepting the offer of acupuncture had been to find an acceptable and effective form of treatment.

Step Two: safe and natural

Moreover, all the interviewees had liked that acupuncture might prove a safe, natural alternative to drugs and surgery; their confidence in its' potential effectiveness coming primarily from a previous successful experience and positive reports of acupuncture in other contexts – as none had heard of it being used for menorrhagia.

Step Three: opportunity cost

However, attending for treatment had involved an opportunity cost, as there were both gains and losses to be weighed and considered. First, a considerable amount of time

had to be invested to attend some 20-sessions over a period of six months - time that had to be carefully planned around family responsibilities, and might have been more usefully employed elsewhere. Second, the needle insertion, and stimulating the needle to produce the 'deqi' sensation, was often unpleasant and sometimes painful. And yet, any discomfort from the needles was momentary and the overall experience a pleasant one. They had liked the feelings of relaxation and well-being that came from receiving acupuncture, and that treatment provided regular time-for-self. They had felt listened to, understood, and cared for, and liked that the time spent attending for treatment had allowed a trusting, supportive therapeutic relationship to develop. They had particularly liked and valued the holistic perspective of acupuncture; that treatment did not focus exclusively upon their menstrual bleed, but instead took into account the full array of psychological and physiological symptoms they experienced; that it had involved their coming to understand the interconnectedness of signs and symptoms, and the impact of lifestyle and emotions upon health; and that they had been advised about ways they could help to bring about and maintain their health.

Step Four: adherence (compliance)

Ultimately, the gains were considered to have outweighed the losses where the health benefits from treatment added significantly to quality of life. Thus, where benefits were less immediate, adherence to the full course of acupuncture had required patience and a steady determination.

Step Five: worth it?

As all the interviewees had experienced some health benefits by six-month follow-up, attending for treatment was considered to have been worth-the-while. Indeed, the fact that their had been health improvements, even if not to the menstrual bleed itself, was considered a most important aspect of treatment acceptability.

Step Six: treatment acceptability

However, whether acupuncture continued to be considered an acceptable treatment option for menorrhagia depended upon the nature of the health gains achieved by the end of the ACUMEN treatment period and six-month follow-up. Those who had found acupuncture effective in normalising their menstrual loss asserted that they would wish to return for another course of acupuncture, should their menstrual

symptoms return. Whereas those who had not experienced relief from symptoms of heavy menstrual loss, wished now to consult their GP for a new drug therapy – even if they had, prior to ACUMEN, considered medication unacceptable because of the risk of potential side effects.

Steps Seven: willingness-to-pay

The health gains experienced during ACUMEN also determined the role the interviewees' perceived for acupuncture in their future health-care; that is, whether they'd use acupuncture for menorrhagia, certain other health problems, or for both. And because they could choose to continue to use acupuncture outside of the trial, and because this option raised cost implications, treatment acceptability came to be expressed in terms of their willingness-to-pay out-of-pocket for acupuncture - and they were willing-to-pay for treatment, where they could be relatively certain that the desired health outcomes would be achieved.

6. Discussion

6.1 Summary of findings

This study involved the analysis of data from interviews with six of the 20 patients randomised to and accepting the offer of acupuncture, in addition to standard GP care, in the York and Selby Acupuncture for Menorrhagia Exploratory trial. The different components of 'treatment acceptability' were explored prior to, during and following a course of treatment, and for the very greater part, the interviewees' lived experiences of acupuncture were positive – even for those who had not found it effective for menorrhagia.

The results from the initial interview illustrated, first, how the sample had considered acupuncture to be both a credible and attractive treatment option for their symptoms of excessive menstrual loss and pain. Indeed, they had been delighted with the opportunity to receive through their GP, what they perceived to be a safe, natural, and 'hands on', low-tech alternative to drugs and surgery - even though prior to ACUMEN, none had heard of acupuncture being used for menorrhagia. Confidence in its' potential effectiveness had come primarily from a previous successful

experience and positive reports of acupuncture in other contexts. And their primary motive for wishing to try acupuncture had been to find an acceptable and effective form of treatment; their collective experience of drug therapy being that it was unavailable, ineffective, inappropriate given their medical history, or unacceptable because of potential side effects. Thus, this sample of patients reported both a need for alternatives to the therapies currently available from their GP for menorrhagia, and an interest in, and preference for, acupuncture as a natural, low-tech option.

The results from the follow-up interviews revealed how the interviewees considered acupuncture to be an acceptable treatment option for menorrhagia. All had felt their referral for acupuncture had been appropriate, and none, with the benefit of hindsight, would have wished to change that decision. Overwhelmingly, these women had found attending for treatment a pleasant, relaxing experience. They had enjoyed a caring, trusting and supportive therapeutic relationship. And they had particularly liked and valued that treatment had not focused exclusively on their menstrual bleed, but had instead taken into consideration their entire experience of health, and provided them with tools by which to attain and maintain health (i.e. dietary and lifestyle advice). These women had liked that they had come to experience a range of health benefits during ACUMEN, including a reduction in menstrual loss and pain, more energy, fewer migraine headaches, resolution of a shoulder problem, and a sense of general well-being. And it was these improvements that had made adhering to treatment and attending for up to 20-sessions of acupuncture over a period of some six months worth the considerable time invested and the discomfort from the needles, regardless of whether or not treatment had succeeded in normalising menstrual loss. However, acupuncture only continued to be considered an acceptable treatment option for menorrhagia by those women who had found it effective in this context. Those who had not experienced relief from symptoms of heavy menstrual loss stated that they would not be willing to embark upon another course of acupuncture, and would instead be consulting their GP for a new drug therapy.

Moreover, the wide range of health benefits experienced meant that the concept 'treatment acceptability' and the role they had come to define for acupuncture in their future health care was not restricted to menorrhagia. Rather, what acupuncture was considered acceptable as a treatment for, whether they'd wish to use acupuncture for

menorrhagia, certain other health problems, or for both, was determined by the health gains experienced in *ACUMEN*. And because they could choose to continue to use acupuncture privately outside of the trial, and because this option raised cost implications, treatment acceptability came to be expressed in terms of their willingness-to-pay out-of-pocket for acupuncture - and they were willing-to-pay for treatment, where they could be relatively certain that the desired health outcomes would be achieved. And for four women who had previously found conventional medicine to be ineffective, inappropriate because of their medical history, or unacceptable because of potential side effects, acupuncture came to be considered an acceptable and effective treatment option for menorrhagia that they were willing to pay for.

6.2 Transferability of findings

Clearly, should these findings be considered transferable to the acupuncture group as a whole and, by extension therefore, to all 39 women participating in this randomised trial, they could be seen to provide some explanation for why patient-centred recruitment strategies proved the most successful (e.g. GP surgery poster and local newspaper feature) *[Chapter Five]*, and why all 20 of the women randomised to acupuncture completed their course of treatment *[Chapter Three]*. Moreover, if the results presented here are considered transferable to women eligible to participate in future trials, these findings could go on to provide useful guidance when designing the trial - particularly with respect to recruitment strategies, information sheets and delivery of the intervention. How confident we can be that these findings are transferable to other women, is guided first by how similar the interviewees appear to be to other women who might participate in future trials, and second, by the level of congruence with prior theory.

The sample was generally similar to the acupuncture group as a whole with regards to the demographic characteristics recorded (i.e. ethnicity, age, education, marital status, number of children, housing, transport arrangements and whether they lived in York or Selby). The three dimensions they differed on were employment, entry route to the study and dispersion amongst the three ACUMEN acupuncture practitioners: Whereby, women engaged in interviews were more likely to be full-time looking after

the home and family and less likely to be employed full-time; they had come to the study via the two most popular routes (GP surgery poster and acupuncturist recommendation) or following an invitation from their GP, but they had not come in response to a feature about ACUMEN in a local newspaper, or letters of invitation from their GP; and most had received treatment from either practitioner B or C, when almost half of the women in the acupuncture group had instead consulted practitioner A.

These differences are worth considering. First, because interviews with more than one woman in full-time employment may, for example, have revealed practical problems affecting treatment acceptability, such as a lack of appointment times outside of office hours - although there was no evidence of this by way of patients' failing to comply with their full course of treatment. With regards the routes by which women came to be involved in ACUMEN, whilst GP letters to eligible patients are not dissimilar in essence to GP surgery posters - both were aimed at women and required that they initiate their own recruitment, and both carried their GP's stamp of approval, because the information had been provided by them - the women who contacted the researcher in response to the newspaper feature could not have assumed their GP's support. And finally, it is possible that interviews with more patients consulting practitioner A would have generated different results, because of the diversity that exists in practice styles (MacPherson & Kaptchuk 1997).

6.3 Congruence with prior theory

One way to further examine these threats to transferability is to ascertain whether the findings are congruent with, or confirm, prior theory (Miles & Huberman 1994; Silverman 2000). For example, other researchers have reported upon how CAM patients are most often motivated by a rational and pragmatic concern to get relief from specific, chronic symptoms, and on terms that are acceptable to them (Barrett et al 2003; Luff & Thomas 2000; Gamon 2000; Paterson & Britten 1999; Cassidy 1998a); that they have often exhausted conventional medical options prior to receiving CAM; that they have wished to avoid or reduce their medication, because of concerns about side effects; and that these views and experiences have occurred alongside an interest in natural, low-tech interventions. These are themes that were

also evident in this study, and would suggest that patients' initial motivations for using CAM are consistent across settings and user groups.

Furthermore, the diverse range of benefits experienced by the women in this study are similar to the findings in four other qualitative studies of traditional acupuncture users in England (Walker et al. 2004; Patterson & Britten 2003; Gould & MacPherson 2001) and the United States (Cassidy 1998), two of which were published after this study had been designed and conducted (Walker et al. 2004; Patterson & Britten 2003): Walker and colleagues undertook focus groups with 16 of the 50 women who had received acupuncture for menopause-like symptoms associated with Tamoxofen in a pilot study. The changes reported included relief from general aches and pains and lethargy. Patterson and Britten interviewed 23 patients with a chronic illness who were having acupuncture for the first time from one of eight practitioners. Each patient was interviewed three times over a period of six months. These patients described improvements to both the presenting symptom and other concurrent or subsequent symptoms, along with increased energy, by way of physical and emotional stamina and resilience. Gould and MacPherson's questionnaire survey of 72 patients attending for treatment with four practitioners revealed that most of the patients who had sought help for a physical problem also experienced improvements in their moods and well-being – "holistic benefits". And Cassidy analysed 460 handwritten stories collected as part of a survey of 575 patients attending for treatment at four private and two teaching clinics. She reported on the "expanded effects of care", whereby patients often experienced relief of presenting symptoms in conjunction with improvements in energy levels, improved immunity and a sense of well-being or wholeness.

Moreover, these findings are congruent with acupuncture theory. Schulman (2004) uses two case studies to illustrate the way that these 'unexpected benefits' of acupuncture do in fact *"follow logically from the central therapeutic imperative of Oriental medicine; treatment of both the patient's root and branch"* – For example, Kidney Yang Deficiency can manifest simultaneously as menorrhagia, fatigue, feeling fearful and low back pain (Maciocia 1998)- and by treating the root – Kidney Yang Deficiency – it then becomes, Schulman asserts, *inevitable* that more than one branch symptom responds to treatment, and rather than being a secondary feature of

acupuncture, these "holistic benefits" or "expanded effects of care" are central to, and evidence of, correct treatment.

Patterson and Britten also described changes in what they termed "personal and social identity", which relates to the way patients gained insight into the causes of symptoms (self-awareness), made lifestyle changes (self-help), reduced their medication usage and learnt to "listen" to their body (self-responsibility or acceptability). This study mirrors these findings within the themes "a holistic perspective" and "perceived outcomes from treatment"; where the value patients placed upon learning to understand the interconnectedness of signs and symptoms and the impact of lifestyle and emotions upon health is described, along with the sorts of lifestyle changes, or shifts in perspectives, that were made as a consequence. These beneficial changes in perspective and to lifestyle were also found by Cassidy and grouped under the category "improved psychosocial coping". Gould and MacPherson observed similar changes, and an improved ability to cope with symptoms that had not changed in response to acupuncture, through an understanding of the nature of their disharmony, was an outcome found by Walker and colleagues.

In this study, dietary and lifestyle advice according to the theories of Oriental medicine that underlie acupuncture, were a common, and often prominent, aspect of patients' care, and for some, being educated in "self-help" was highly valued. The giving and receiving of lifestyle advice was alluded to, or noted, by two of the previous studies (Walker et al. 2004; Gould & MacPherson 2001), and reported as an integral aspect of treatment by the others (Patterson & Britten 2003; Cassidy 1998b). Patterson and Britten looked in particular at the way in which advice was given, using descriptors such as "suggested", "supported" and "encouraged," and noted that by contrast, "directive advice" played only a minor role in treatment. Similarly, Cassidy (1998b) described how patients were "empowered" to bring about lifestyle changes in response to what she termed "education and encouragement," and determined that this outcome from treatment was entirely dependent upon the close, facilitative, or horizontal, patient-practitioner relationships patients' reported. Moreover, Patterson and Britten's analysis led them to theorise that not only is a close, respectful relationship the means by which lifestyle advice is successfully delivered by acupuncturists, but that the ability to deliver pertinent advice in a supportive and non-

judgemental way, further strengthens the relationship through a process of cyclical reinforcement. This study can be seen to add to the mounting evidence that holistic principles often engender caring, trusting and supportive therapeutic partnerships (Walker et al. 2004; Barrett et al 2003; Gould & MacPherson 2001; Luff & Thomas 2000; Paterson & Britten 1999).

The results of this study reflect recent research findings about patients' prior anxieties about needles and their experience of insertion and manipulation as often unpleasant and occasionally painful (Walker et al. 2004; Patterson and Britten 2003). They also reflect the finding that any discomfort is momentary and the overall experience a pleasant one (Walker et al. 2004; Cassidy 1998b) that often gives a profound sense of relaxation and calm that might last for sometime after treatment (Walker et al. 2004; Patterson and Britten 2003; Cassidy 1998b). However, none of these studies found the time involved in attending for treatment to be a component of the treatment process that detracted from patients' experiences of care. The role of regular visits in the development of a trusting, therapeutic partnership was recognised, but the opportunity cost of time in terms of whether the time spent in treatment might not be better spent elsewhere, or difficulties in finding the time for treatment around work and family commitments, have not previously been reported. It is possible that this finding is related first, to the fact that menorrhagia requires treatment over a minimum of two to three cycles, meaning that condition specific benefits may not have been experienced in the first 10-sessions of acupuncture, so affecting motivation; second, to the fact that the trial imposed the condition, or expectation, that patients would attend for up-to 20-sessions of acupuncture, over a period of some six months, which may have appeared daunting; and third, to the fact that the sample necessarily comprised women with family commitments, half of whom also had work commitments. And whilst patients who attend for acupuncture often have a chronic health problem and may come to receive in excess of 20-sessions, there is no evidence to suggest that this was ever expected, or assumed, by either the patient or the practitioner (Patterson & Britten 2003; Gould & MacPherson 2001; Cassidy 1998). And although the women in the pilot study of acupuncture for the menopausal-like symptoms associated with Tamoxofen were asked to commit to a course of treatment, it consisted of just eight sessions. Thus this finding has implications for other trials of acupuncture that

involve lengthy treatment courses, and further qualitative studies that are nested within acupuncture trials might usefully test this hypothesis.

6.4 What this study adds

Ultimately, by interviewing a specific patient population prior to, during and following a planned and funded course of acupuncture, this study has provided unique insights into the way in which what patients like and value – the components of "treatment acceptability" – might evolve over time, in line with lived experiences.

In this study, what initially determined treatment acceptability, and therefore, credibility, was the perception of acupuncture as a safe and natural alternative to drugs and surgery, alongside positive experiences, or reports, of acupuncture – but in other contexts, as none had heard of acupuncture being used to treat menorrhagia.

Once treatment began, acceptability came to be determined by the complex interplay of components related to 'the acupuncture treatment *process'*, 'the holistic *perspective* of treatment' and 'the perceived *outcomes* from acupuncture' *[thematic framework, Table 3]*:

a. Process: clinic environment, therapeutic relationship, sensation from the needles (i.e. pain and relaxation), time required (i.e. regular time for self versus difficulties juggling work and family commitments to accommodate treatment),
b. Perspective: treatment took into consideration full array of physiological and psychological signs and symptoms, learning to understand self as a whole, advice give to attain and maintain health,
c. Outcomes: normalisation of menstrual loss, improvements in other symptoms and well-being, positive changes in health behaviour.

Whereby those aspects of *process* and *perspective*, in conjunction with, and more importantly mediated by *outcomes*, determined treatment acceptability and patients' willingness to adhere to the full-course of treatment. This interplay was expressed in terms of the opportunity cost involved in attending for treatment, where patients considered the gains (e.g. caring and supportive therapeutic relationship, learning to

see self as a whole, and the relaxation and calm induced) to outweigh losses (time required and pain on needle insertion) where health benefits added significantly to quality of life.

Thus these findings are congruent with Patterson and Britten's (2004) finding that positive effects from treatment work to reinforce patients' confidence in acupuncture and its' holistic basis. Their description of the complex interplay and evolution of the components concerned with 'process' (therapeutic relationship, diagnostic and needling skills, new holistic understandings) and 'effects' (changes in energy, strength and relaxation, changes in symptoms, changes in personal and social identity) are all evident in this study, but are here couched in terms of patients' likes and dislikes and encapsulated by the themes *process, perspective* and *outcomes.*

Come the end of treatment, treatment acceptability, or more precisely, what role patients predicted acupuncture would play in their future health care, was determined by the nature of the benefits experienced – and a second course of acupuncture was not considered acceptable to the women whose symptoms of menorrhagia had not responded to treatment. Therefore the results from this study further illustrate the finding that patients are discerning consumers of CAM and conventional medicine (Patterson & Britten 2003; Gould & MacPherson 2001; Luff & Thomas 2000; Patterson & Britten 1999; Cassidy 1998), whereby the same pragmatism that led women to try acupuncture for menorrhagia, when it was found to be ineffective, led them later to return to their GP for a new drug therapy. And conversely, where acupuncture successfully addressed the main complaint, concurrent symptoms and well-being, the role perceived for acupuncture expanded beyond the initial remit of menorrhagia.

This study also adds to the evidence that after successful treatment of the original problem, patients' use of acupuncture may move toward maintaining health and well-being (Patterson & Britten 2003; Gould & MacPherson 2001; Cassidy 1998). Moreover, it adds to the evidence that commitment to treatment for many NHS patients develops as a direct result of satisfaction with treatment, and not because of any commitment arising from financial investment (Luff & Thomas 2000), but suggests that for NHS patients receiving a course of fully-funded treatment as part of

a trial, satisfaction with treatment can result in financial investment in private acupuncture – a willingness-to-pay out-of-pocket - in order to maintain the benefits experienced, or in order to access what they perceive to be the most appropriate therapy for a given health problem, based upon their lived experience. This is of interest, particularly with respect to the maintenance of benefits following successful treatment of the condition under investigation, as this is not currently a feature of acupuncture effectiveness studies; even though it has implications in terms of optimum treatment in to the long-term, and the real costs of the intervention and where that burden falls.

An interesting difference between this study and those conducted by Patterson and Britten (2003 and 1999) and Luff and Thomas (2000) concerns the way in which data about therapeutic relationships was organised. These studies recognised the way in which patients often contrasted the relationship they had with their CAM practitioner with those they had with their GP, and so generated a theme or sub-theme for the lack or loss of caring, therapeutic GP relationships. This was not done here, although there is evidence of this tendency to compare and contrast, and would seem to accurately reflect a limitation of this study – namely that the candidate is also an acupuncturist, which in this instance led her to deflect from this finding, in order to avoid a seemingly positive bias toward acupuncture and acupuncturists and suggest a hostility toward GPs.

6.5 Planned future research

The interviews conducted for the purposes of this study of 'treatment acceptability' took place within in a week of administering the baseline and posting the three and six-month trial questionnaires. These questionnaires included validated, health outcomes measures: the SF-36 and Aberdeen Menorrhagia Questionnaire, alongside measures developed for the purposes of ACUMEN in order to ascertain the wider effects of treatment upon menstrual health and to provide global markers of change of immediate relevance to patients *[Chapter Two]*. The candidate aims to compile individual health outcomes profiles using the interview data and questionnaire data. The scores from the measures will provide a context for patients' descriptions of change and thereby increase the comprehensiveness of the findings via the

triangulation of different data types. This analysis will also allow the candidate to gauge the extent to which these instruments encompassed and measured the full-range of treatment effects patients reported at interview, with the aim of informing future trials seeking to evaluate the effectiveness of acupuncture for women with menorrhagia in primary care. For whilst outcome measures cannot be expected to capture all aspects of change, the evidence from previous research is that if carefully chosen, it should be possible to register most of the important treatment effects most of the time (Patterson and Britten 2004; Long et al. 2000; Patterson & Britten 2000; Tully & Cantrill 2000; Hickey et al. 1996; Hill et al. 1996; Ruta et al. 1995).

6.6 Implications for future research

Despite the small sample size, the methodological rigour and degree of congruence with previous work would support using these findings to inform future research, particularly with regard to the design and conduct of future trials of acupuncture for menorrhagia and the development of recruitment strategies, information sheets and delivery of the intervention.

First, these findings lend support to the use of patient-centred recruitment strategies in future trials, as it would appear that these succeeded by tapping into a need for alternatives to the therapies currently available from GPs for menorrhagia, and an interest in, and preference for, acupuncture as a natural, low-tech option. Moreover, it is likely that future research will be able to enhance the effectiveness of methods such as GP surgery posters, letters from GPs and newspaper features, by reporting the findings from the ACUMEN trial and describing the treatment outcomes for the 6 women who participated in interview.

These results would also suggest that potential barriers to participation, namely fears about needles and concerns about the amount of time that would be involved in attending for up to 20-sessions of acupuncture, be attended to from the outset in the information sheet. For example, vignettes compiled from the data collected for this study would both illustrate, and provide a context for, information about needle sensations, making it more accessible and better able to allay fears. The positive and negative aspects of regularly giving time to attend for treatment should be set out.

And most importantly, the information sheet should describe traditional acupuncture's holistic approach and the range of benefits that patients often experience as a consequence and which can make attending for acupuncture "worth it" - even where normalisation of menstrual loss is not experienced; again, vignettes or 'stories' compiled using the data from this study could be usefully employed here.

Future trials should also look to minimise the time spent travelling to treatment by providing as accessible an acupuncture service as possible for the patient population involved – it is worth noting that for some patients, participating in ACUMEN involved their travelling some 15-miles for treatment, and that one of the three practitioners did not provide treatment outside of office hours (i.e. after 5pm). However, the importance of the clinic environment and where these acupuncture services might best be located in future trials is less clear from this study, as only one patient made specific reference to the way in which the treatment setting had impacted upon their experience of acupuncture. Thus, researchers are advised to await the publication promised by Patterson and Britten (2004) about the impact of the institutional context of the intervention upon experience and outcome, before deciding whether the problem of access might be successfully addressed by providing acupuncture at both private clinics, as in ACUMEN, and also at GP surgeries for the purposes of the trial.

Future qualitative studies might usefully test the validity of the model, developed to describe the way in which 'treatment acceptability' evolved over time in this study sample, for both a larger sample of women receiving acupuncture for menorrhagia within a trial, and for patients receiving acupuncture for other conditions. Researchers might also assess the relevance of the opportunity cost of time to treatment acceptability in other trials of acupuncture, and particularly in those that involve lengthy treatment courses. The implications that arise from patients privately accessing maintenance treatment following a successful course of acupuncture for the condition under investigation in a trial should also be considered; especially with respect to what might constitute optimum treatment into the long term, the real costs of the intervention and where that burden falls.

Table 42a: Quality issues in qualitative research, from *"An expanded sourcebook: qualitative data analysis"* by Miles & Huberman 1994, pages 277-280

A. Objectivity / Confirmability	1. Explicit methods 2. Conclusions linked with displayed data 3. Self aware about biases and assumptions 4. Competing / rival hypotheses considered 5. Data available for re-analysis by others
B. Reliability	1. Clear question and congruent study design 2. Researcher's role and status explicit 3. Connectedness to theory 4. Coding checks 5. Colleague review
C. Credibility / Internal validity	1. Context rich thick descriptions 2. Triangulation gave generally converging conclusions 3. Categories of emerging theory clearly linked to the data 4. Coherence 5. Areas of uncertainty identified 6. Negative evidence sought for 7. Rival explanations considered
D. Transferability / External validity	1. Characteristics of sample fully described 2. Threats transferability examined 3. Sampling theoretically diverse enough to encourage broader applicability 4. Congruent with or confirms prior theory 5. Narrative sequences are unobscured and cross case theory using sequences is developed 6. How to do further testing suggested

Table 42b: Quality issues in qualitative research, from *"Doing qualitative research: a practical handbook"* by Silverman 2000, pages 175-190

Validity is achieved via the "refutable principle" and involves 4 inter-related methods:	
A. Constant comparative method	1. Provisional hypotheses tested on subsequent cases
B. Comprehensive data treatment	1. Dataset is open to repeated inspection until generalisation is able to apply to every single gobbet of relevant information 2. Integrated, precise model that comprehensively describes the phenomenon is produced
C. Deviant case analysis	1. Provisional analytic scheme constantly confronted by discrepant cases until small set of recursive rules that incorporate all data in analysis is derived
D. Tabulation	2. Theoretically derived counting techniques used to test and revise generalisations (where appropriate)
Reliability is dependent upon the procedure – the evidence of reliable methods	
***Generalisability* or extrapolation is achieved when there is:**	
	1. Purposive sampling 2. Theory driven analysis 3. Congruence with other work

6.7 Strengths and weaknesses of the study

This study was designed and conducted with attention to many of the criteria suggested for increasing validity in qualitative research by Miles & Hubberman (1994) and Silverman (2000) *[Table 42a & 42b]*. The design, conduct and analysis followed the methodological frameworks of both descriptive analysis and framework analysis and are described in detail (Sandelowski 2000, Ritchie & Spencer 1994). The candidate's role and status is made explicit. Purposive, reflexive sampling strategies were adopted to maximise the transferability of the findings, and the characteristics of the sample are described. The interviews involved prolonged engagement (45-90 minute interviews) and repeated contacts. The development and refinement of the thematic index took place at the same time as coding the transcripts, and involved a to-ing and fro-ing between the coded transcripts and index in order to continually test assertions made about the relevance of themes and the connections between ideas. The case charts and thematic charts generated work to better facilitate access to the patient's own words and allow the analytical process to be examined and replicated. And whilst the themes were not verified independently, the emergent index, or framework, was agreed by the candidate's co-supervisor (KT). The findings are qualified and illustrated with referenced data. The process of interpretation, where the patterns and connections between ideas, views and experiences are mapped out, is clearly linked to the data. The congruence of the findings to other research is demonstrated, and threats to transferability are examined.

However, these strengths, including a self-awareness about biases and assumptions, cannot detract from the key weakness of this study, which is its' small sample size. Samples of 20 or more are usual for this sort of applied policy research, where specific information needs culminate in actionable outcomes (Murphy & Dingwall 2001). A sample size of six lessens the transferability of the findings, even where there is evidence of reliable methods, the characteristics of the sample are largely equivalent to the sample frame, and the congruence of findings with other research is evident. Less frequent interviews, for example, omitting the three-month follow-up and instead conducting interviews at baseline and six-months follow-up, would have allowed for a sample size of *10*, so comprising half the acupuncture treatment group. The finding that patients in this study tended to speak of their entire treatment

experience, and that the views and experiences they gave at three months were congruent with those given at six months, would serve to support less frequent interviews. Although arguably, it was the three-month interview itself that ensured an equally reliable account of their early experiences at six-month follow-up, and omitting it would have given a less comprehensive account.

6.8 Conclusions

This longitudinal qualitative study provides evidence of the components of treatment acceptability and how these evolved over time in line with lived experience for six women with menorrhagia in primary care receiving a fully funded course of acupuncture in the ACUMEN exploratory trial. Despite the small sample size, the methodological rigour and degree of congruence with previous work would support using these findings to help explain the success of patient-centred recruitment strategies in ACUMEN and the 100% compliance to acupuncture. These findings also offer valuable guidance for the design and content of recruitment material, the content of patient information sheets and delivery of the complex traditional acupuncture intervention in future trials of acupuncture for menorrhagia in primary care.

Chapter Five

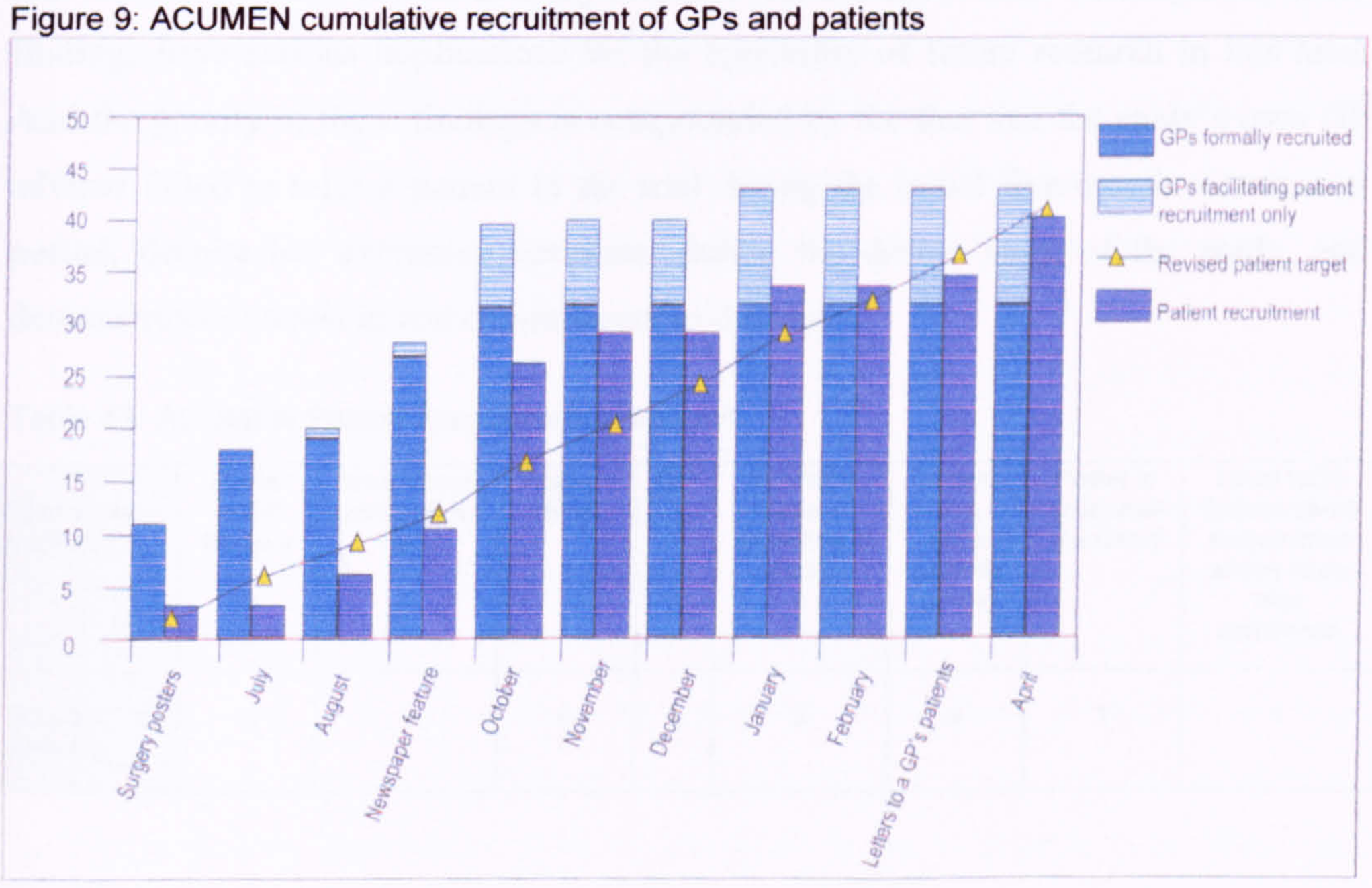
The ACUMEN GP Survey:

A survey to identify barriers to active GP participation in an exploratory pragmatic randomised trial of acupuncture for menorrhagia

1. Survey Rationale

The York and Selby Acupuncture for Menorrhagia (ACUMEN) Exploratory Trial was the first UK-based study of acupuncture in the treatment of menorrhagia in primary care (Longridge et al. 2001). The trial involved 39 of a target of 40 patients. The principle aim of the trial was to inform future research as to the justification for, and feasibility of, a full-scale pragmatic randomised trial to evaluate the clinical and economic benefits of offering acupuncture to patients with menorrhagia in primary care. Thus the ACUMEN trial was designed to answer questions about the relative clinical and cost effectiveness of different treatment options for menorrhagia, including acupuncture, in the real world setting of primary care. And as a consequence, the success or failure of the ACUMEN study was largely dependant upon the willingness of general practitioners (GPs) to refer suitable patients to the trial.

Figure 9: ACUMEN cumulative recruitment of GPs and patients



The ACUMEN research team predicted that the recruitment target of 40 patients would be achieved over the course of five months with the support of 10 GPs. Introductory letters, enclosing a study summary and reply postcard for those interested, generated a positive response from 11 GPs, who then enabled the enrolment of a further eight GPs from their practice. However, after three months just six patients had been recruited to the study compared with a target of 24. As a result the recruitment period was extended by six months and the researcher continued to ask for GP support. After five months, 32 GPs (from 15 practices) in the Selby and York Primary Care Trust had formally offered their support. And by the end of the recruitment period, another 11 GPs had facilitated recruitment for 16 patients who first contacted the researcher to express an interest in taking part. Thus a total of 43 GPs were involved in patient recruitment *[Figure 9]*. Although, of the 39 patients recruited out of a target of 40, only four were GP initiated *[Table 43]*. The majority of referrals were initiated by the patient in response to patient-centred recruitment methods and facilitated by the GP. Posters in GP surgeries and a local newspaper article about the study were the most effective recruitment methods. It is notable therefore, that almost half of ACUMEN patients (18) came into the study after first contacting the researcher. But what these findings reveal in particular, is a large discrepancy between the number of GPs who formally agreed to support the study (32/43) and the number who actually initiated recruitment (4/32). Consequently these findings have serious implications for the feasibility of future research in this area. And the gravity of these findings is compounded by the fact that the study's own GP advisor failed to refer a patient to the trial during the initial five-month recruitment period, despite her expressed optimism during the design stage of the study, and demonstrated interest in and commitment to the study.

Table 43: ACUMEN Patient Recruitment Pathways

Recruitment pathway	Poster in GP Surgery	Newspaper article about study	Acupun-cturist	GP	ACUMEN patient recommend-dation	Letters to a GP's patients (n=14)	Poster in Acupunct-ure Clinic	Local radio feature about acupuncture where study was mentioned
Number of patients	12	9	6	4	2	4	1	1

It is not uncommon for trials to be compromised, or even fail, on account of failing to obtain sufficient recruitment rates (Fairhurst & Dowrick 1996; Hunt et al. 2001;Peto et al. 1993; Tognoni et al. 1991; Van der Vint et al. 2000). The literature covering problems of recruitment to trials describes a diverse range of factors that have been found to militate against maximal recruitment, including the time pressures faced by GPs within the consultation, their uncertainty as to the benefits of the experimental intervention, and a complex trial protocol (Bell-Syer & Klaber Moffett 2000; Prescott et al. 1999; Ross et al. 1999). As GPs are most familiar with acupuncture as a treatment option for conditions characterised by pain (British Medical Association 2000), uncertainty as to the benefits of acupuncture for menorrhagia may well be the key factor. It is notable that 43 GPs successfully recruited 241 patients over a period of 18 months for the York Acupuncture for Back Pain Study (YACBAC) (Thorpe & Bell-Syer 2002). As the YACBAC and ACUMEN randomised trials share much in common, from the pragmatic design, members of the research team, GP population, and acupuncture service provided (Thomas et al. 1999), the key differential variable is arguably the condition treated – back pain being both more common than menorrhagia and, perhaps more importantly, a condition characterised by pain; a hypothesis that has implications for the field of acupuncture research generally.

It is significant therefore, that the study's GP advisor, having referred just one patient to the study after eight months of recruitment, agreed to trial a database recruitment method. This involved identifying patients listed as having consulted for treatment for idiopathic menorrhagia within the past six months, and yielded 14 eligible patients who had not been invited to participate during their consultation. The considerable time the GP advisor spent developing the letter of invitation to patients and accompanying information about the trial, screening the lists of patients their practice manager had generated by searching the computerised practice list, and providing the researcher with the relevant contact details, further demonstrated their commitment to the study. It also suggested that this might prove a more acceptable, and therefore efficient, recruitment strategy. It is notable that a recent study of acupuncture for chronic headache in primary care successfully recruited 401 patients by asking GPs to send letters and trial information to potential participants identified on their computerised practice lists (McCarney et al. 2002; Vickers et al. 1999). Furthermore, such retrospective recruitment methods are likely to be acceptable to patients with

menorrhagia, as the four patients recruited to the ACUMEN trial in this way reported feeling first grateful to their GP for remembering them, and second privileged to be given the opportunity to take part in the trial.

Thus a survey designed to identify the barriers to active patient recruitment by GPs in the ACUMEN trial was considered essential at this planning or exploratory stage, in order to assess the degree to which these barriers might be overcome and more efficient recruitment strategies developed. This latter point is significant, as the costs of recruitment are a major portion of the budget for many full-scale trials, especially where recruitment periods have to be extended to ensure the statistical power of the study is not decreased (Ross et al. 1999). And as a need to extend the recruitment period is increasingly being used as a "quality indicator", studies may not attain either the funds or ethical approval necessary to achieve their recruitment targets in this way (Prescott et al. 1999).

2. Aims of the ACUMEN GP survey

I. To identify barriers to active GP participation in the York and Selby acupuncture for menorrhagia exploratory trial,
II. Assess the degree to which these barriers might be overcome and more efficient recruitment strategies developed, and in so doing
III. Inform future research as to the feasibility of a full-scale trial that relies upon GPs to initiate patient recruitment.

3. Objectives

I. To identify factors that increased GP willingness to refer patients to the ACUMEN trial,
II. To identify factors that decreased GP willingness to refer patients to the ACUMEN trial,
III. To test the hypothesis that the key barrier to active GP participation in the ACUMEN trial was GPs unfamiliarity with acupuncture as a treatment option for conditions other than those characterised by pain,
IV. To assess the acceptability of an alternative recruitment pathway involving GPs writing to suitable patients listed on the computerised practice list,

v. To seek GPs views of other ways of facilitating recruitment to trials like ACUMEN.

4. Methods

4.1 *Questionnaire development*

Questionnaire development was guided and informed by the following:

a) A review of the relevant literature and,
b) The York Acupuncture for Back Pain GP survey.

Additional questions were included:

c) Based upon factors reported by GPs during the ACUMEN trial and recorded in the researcher's field diary, and
d) Designed to test the hypothesis that the key barrier to active GP participation in the ACUMEN trial was GPs unfamiliarity with acupuncture as a treatment option for conditions other than those characterised by pain, and
e) Designed to assess the acceptability of an alternative recruitment pathway and seek GP views of other ways to facilitate patient referral to trials like ACUMEN.

In this way, the questionnaire sought to satisfy the key requirement for content validity that it is comprehensive. This process is described in 4.1.1 to 4.1.5 below:

4.1.1 Questionnaire development: A review of the relevant literature

A MEDLINE search using the keywords recruitment, primary care, RCT and referral yielded 14 papers dated between 1989 and 2001 that noted, or focused specifically upon, factors that had helped or hindered GP participation and subsequent patient referrals in trials of healthcare interventions (Bell-Syer & Klaber Moffett 2000; Borgiel et al. 1989; Foy et al. 1998; Hunt et al. 2001; Jonker & Sumajow 1992; Lovato et al. 1997; Murphy et al. 1992; Peto et al. 1993; Pringle & Churchill 1995; Ross et al. 1999; Silagy & Carson 1989; Tognoni et al. 1991; Van der Vint et al. 2000; Ward et al. 1999).

Table 44: Factors that hinder and help GP participation and patient referral in trials *(page 1 of 2)*

	Sub-heading	Factors that Hinder	Paper(s) citing factor *		Factors that Help	Paper(s) citing factor *	
1	**GP recruit information**				a. Study summary provided prior to personal contact to gain GP co-operation b. Attractive and accessible information pack c. Information pack included photocopy of published feasibility study d. Information pack included policy statement about confidentiality e. Flowchart for the wall of each office provided	a.2,10 b.10 c.2 d.2 e.1	
2	**Research question**	a. Insufficiently interesting question b. See few eligible patients	a.4 b.5	2	a. Important research question b. Personal interest in research	a. 1,3,10 b. 4,7,9,11	7
3	**Intervention**	a. Unsure of benefits of experimental intervention b. Concern for patient as would feel responsible if did not receive what was later found to be the "best" treatment c. Cautious about use of non-drug treatments because of the difficulty of ensuring quality control d. Belief that patient would benefit more from another intervention	a.1,6 b.4 c.6 d.3	4	a. Perception of patients receiving superior clinical care b. Lack the service the trial will provide and like being able to offer it to patients	a.7,10 b.1	3
4	**Time**	a. Time constraints within consultation b. Excessive time commitment for either GP or patient c. Already involved in too many research projects	a.1-4 b.5-7 c.1,5	7	Demands on participants (GP and patient) kept to a minimum	3,4,8,9 10	5
5	**Protocol**	a. GP obtaining informed consent is problematic b. Protocol too complex c. Complexity of forms d. Uncertain of entry criteria e. Protocol too difficult to comply with – goes against normal practice f. Disapprove of / dislike trial design g. Loss of professional autonomy	a.4,6,7,12 b.13 c.13 d.5 e.2,6 f.7 g 4	7	a. Trial design incorporates patient preferences b. Protocol is clear and straightforward c. Simple entry criteria (researcher applies detailed inclusion/exclusion criteria) d. Simple and short referral forms e. Researcher obtains informed consent f. Research assistants collect data	a.4 b.4,8,9 c.11 d.4,14 e.1,3 f. 9	7
6	**Patient recruitment**				a. Maximise number of GPs involved b. Use a small number of motivated GPs c. Use nurse practitioners to tell patients about study d. One month on, one month off recruitment to avoid fatigue	a.11 b.3 c.11 d.1 e.11	

					e. Use computerised practice lists to identify suitable patients		
7	**Contact and communication**	a. Lack of recognition for role played b. Lack of adequate research training	a. 1,2,12 b. 2		a. Dedicated, experienced and supportive research team recruit GPs b. Efficient trial management c. Research team forges strong relationships with clinicians (and patients) d. Regular contact with GPs and practice manager e. Regular progress reports f. Letters to GPs about patient involvement g. "Local champion" study GP motivates other GPs h. Personal approach from researchers to invite participation i. Engage GPs from outset	a.4,8,10 b.3,7 c.1,7,11 d.1,3,7,14 e.3,14 f.3,7 g.1,2,9,10 h.9 i.11	10
8	**Forgetfulness**	Forgot!	3,5	2			
9	**Doctor-patient relationship**	a. Randomisation seen as admitting to not knowing which treatment is best for their patient, compromising relationship b. Involvement in experimental studies disrupts doctor-patient relationship	a.4,6 b.12	3			
10	**Practical business issues**	a. Lack of financial compensation / incentives b. Disruption of office routine c. Potential for lawsuits	a.4,7 b.4 c.4	5	a. Provide financial incentives b. Avoid financial incentives – induces scientific fraud c. Provide financial compensation	a.1,6,9 b.14 c.1,3,12,14	6

Table 44 References: #1 Ward E et al. *Conducting randomised trials in general practice: methodological and practical issues,* 49 British Journal of General Practice. 919-922 (1999) **#2**.Borgiel A et al. *Recruiting family physicians as participants in research,* 6 Family Practice. 168-172 (1989) **#3**. Van der Vint D et al. *Practical aspects of conducting a pragmatic randomised trial in primary care: patient recruitment and outcome assessment,* 50 British Journal of General Practice. 371-374 (2000) **#4.** Ross S et al. *Barriers to participation in randomised controlled trials: a systematic review,* 52 Journal of Clinical Epidemiology. 1143-1156 (1999) **#5.** Peto V et al. *Factors affecting general practitioners' recruitment of patients into a prospective study,* 10 Family Practice. 207-211 (1993) **#6.** Hunt C J et al. *Do doctors know best? Comments on a failed trial,* 174 MJA. 144-146 (2001) **#7.** Lovato L et al. *Recruitment for controlled clinical trials: literature summary and annotated bibliography,* 18 Controlled Clinical Trials. 328-357 (1997) **#8.** Silagy C et al. *Factors affecting the level of interest and activity in primary care research among general practitioners,* 6 Family Practice. 173-176 (1989) **#9.** Foy R et al. *Clinical trials in primary care: targeted payments for trials might help improve recruitment and quality,* 317 British Medical Journal. 1169(1998) **#10.** Murphy E et al. *Will you help me with my research? Gaining access to primary care settings and subjects,* 42 British Journal of General Practice. 162-165 (1992) **#11.** Bell-Syer S et al. *Recruiting patients to randomised trials in primary care: principles and case study,* 17 Family Practice. 187-191 (2000) **#12.** Pringle M et al. *Editorial: Randomised controlled trials in general practice,* 311 British Medical Journal. 1382-1383 (1995)**#13.** Tognoni G et al. *Randomised clinical trials in general practice: lessons from a failure,* 303 British Medical Journal. 961-971 (1991) **#14.** Jonker P et al. *Randomised clinical trials in general practice,* 304 British Medical Journal. 508(1992)

The 24 factors thought to have hindered and the 32 factors cited as having helped GP participation and patient referrals are given in Table 44 and grouped under the following sub-headings:

1. GP recruitment information
2. Research question
3. Intervention
4. Time
5. Protocol
6. Patient recruitment
7. Contact and communication
8. Forgetfulness
9. Doctor-patient relationship
10. Practical business issues

Attractive, accessible and concise information, a study summary and office wall chart, were cited as useful ways to engage GP interest in trials in three of the 14 papers (Borgiel et al. 1989; Murphy et al. 1992; Ward et al. 1999). The merit in being able to provide GPs with a copy of a published feasibility study was also noted (Borgiel et al. 1989).

Seven papers highlighted the need to demonstrate the importance of the research question and engage GPs with a personal interest in the area of research (Bell-Syer & Klaber Moffett 2000; Foy et al. 1998; Lovato et al. 1997; Murphy et al. 1992; Ross et al. 1999; Van der Vint et al. 2000; Ward et al. 1999). A lack of interest in the research, or seeing few eligible patients, was cited by two of the papers as a reason for poor GP participation (Peto et al. 1993; Ross et al. 1999).

GP perception of the trial intervention as providing patients with superior clinical care, or an attractive alternative to normal management, was cited as a clear incentive for active participation in three papers (Lovato et al. 1997; Murphy et al. 1992; Ward et al. 1999), while a lack of confidence in the trial intervention, or greater confidence in an existing treatment, was considered a deterrent by four (Hunt et al. 2001; Ross et al. 1999; Van der Vint et al. 2000; Ward et al. 1999).

A lack of time, due to other commitments and current research projects, or due to the study requiring an excessive amount of time from either the GP or patient, was a commonly reported factor that hindered trial participation (Borgiel et al. 1989; Hunt et al. 2001; Lovato et al. 1997; Peto et al. 1993; Ross et al. 1999; Van der Vint et al. 2000; Ward et al. 1999).

Half of the papers cited problems with the trial protocol as reasons for poor GP participation; it being overly complex; requiring that GPs comply with procedures that differ to normal practice, or because compliance would deny them professional autonomy (Borgiel et al. 2001; Lovato et al. 1997; Peto et al. 1993; Pringle et al. 1995; Ross et al. 1999; Tognoni et al. 1991).

To enhance GP participation and patient referral, all the papers advised that trials involve a clear and easily adopted protocol, with succinct referral forms that minimise the time demands on GPs and instead place the burden of paperwork upon the researcher (Bell-Syer & Klaber Moffett 2000; Borgiel et al. 1989; Foy et al. 1998; Hunt et al. 2001; Jonker et al. 1992; Lovato et al. 1997; Murphy et al. 1992; Peto et al. 1993; Pringle & Churchill 1995; Ross et al. 1999; Silagy & Carson 1989; Tognoni et al. 1991; Van der Vint et al. 2000; Ward et al. 1999). However, the 3 papers that advised upon patient recruitment strategies specifically gave differing, although arguably inter-dependent, advice (Bell-Syer S & Klaber Moffett J 2000; Van der Vint D et al. 2000; Ward E et al. 1999). For, whilst two of the papers suggested methods to further minimise the burden upon GPs, namely maximising the number of GPs involved; asking nurse practitioners to inform patients about the study; using computerized practice lists to identify patients; and using a one-month-on, one-month-off policy to avoid fatigue, the third advised that a small, select group of GPs be engaged in the study from the outset to ensure their being well-informed and well-motivated to actively engage in patient recruitment. The reason for arguing that these methods are inter-dependent, as opposed to contrary or mutually exclusive, is based upon the recent experience of researchers involved in a current trial where the initial strategy of engaging a large number of GPs in patient recruitment ultimately facilitated the identification of a small select and particularly motivated group of GPs (Russell 2004).

Ways of motivating GPs were discussed by ten of the 14 papers (Bell-Syer & Klaber Moffett 2000; Borgiel et al. 1989; Foy et al. 1998; Jonker & Sumajow 1992; Lovato et al. 1997; Murphy et al. 1998; Ross et al. 1999; Silagy & Carson 1989; Van der Vint et al. 2000; Ward et al. 1999), the majority advising that researchers look to forge strong working relationships with participant GPs. The merits of engaging a "local champion" GP to gain the support of other GPs was discussed in four papers. And once recruited, the need to keep GPs up-to-date with trial progress was highlighted in three.

It may be that strategies to enhance contact and communication between trialists and GPs would avoid poor participation as a result of "forgetfulness" (Peto et al. 1993; Van der Vint et al. 2000). Although arguably, "forgetfulness" could be a proxy response for factors that are not to do with trial management, and may instead be related to problems with the protocol, or concerns about the trials disruptive impact upon the doctor-patient relationship (Pringle & Churchill 1995).

Two of the papers gave, as reasons for not recruiting patients, GP concern that patients would interpret an invitation to take part in trial as an admittance to not knowing which treatment would be best for them (Hunt et al. 2001; Ross et al. 1999).

Two papers gave practical business issues as reasons for GPs not supporting a trial (Lovato et al. 1997; Ross et al. 1999): a lack of financial compensation, disruption of the office routine and the potential for lawsuits against them, should a trial patient of theirs have a negative experience. To overcome financial barriers, three papers recommended offering financial incentives (Foy et al. 1998; Hunt et al. 2001; Ward et al. 1999). Conversely, one paper advised that financial incentives be avoided, as they may induce scientific fraud, and that instead GPs be given financial compensation, or a fair reimbursement for resources used (Jonker & Sumajow 1992). This was a recommendation reiterated by a further three papers (Pringle & Churchill 1995; Van der Vint et al. 2000; Ward et al. 1999).

4.1.2 Questionnaire development: The YACBAC GP Questionnaire Survey

To facilitate a comparison of findings from a similar questionnaire survey conducted as part of the YACBAC trial (Thorpe & Thomas 2002), the researcher took into consideration the wording used for relevant survey questions: Namely those concerning factors that helped or hindered GP participation, the user-friendliness of the protocol, and the success of strategies undertaken to maintain a high trial profile *[Appendix 50]*. It is hoped that a comparison of GP responses to these questions will enable further testing of the hypothesis that the key barrier to active GP participation in the ACUMEN trial was GPs unfamiliarity with acupuncture as a treatment option for conditions other than those characterised by pain.

4.1.3 Questionnaire development: Factors reported by GPs during the ACUMEN trial

The researcher encouraged GPs to express their thoughts about the ACUMEN trial during meetings to gain their support, or when contacted by telephone. These views were recorded in the researcher's field diary and are summarised under the headings "facilitators" and "barriers" in Table 45 below:

Table 45: Barriers to and facilitators of participation in ACUMEN reported by GPs

	Facilitators	Barriers
1	Wish to support trials of acupuncture	Uncertain of the benefits of acupuncture
2	ACUMEN GP trial advisor is a respected colleague	Uncertain of the benefits of acupuncture for menorrhagia
3	Positive feedback from patients receiving acupuncture in the YACBAC trial	Confidence in the MIRENA coil or COCP
4	Personal positive experience of acupuncture	Uncertain as to whether acupuncture would be acceptable to patients
5	Patient interest in acupuncture	Difficulty for patients attending acupuncture clinics
6	Candidate's supervisor and Head of Department of Health Sciences, University of York, is a respected researcher	An A5 sized desktop information and patient recruitment pack would make remembering to invite patients to participate easier than the A4 folder provided, which can only be stored on a bookshelf or in a drawer
7		Dislike pragmatic trial design, as does not enable a distinction between placebo effects and the 'real' effects of the acupuncture intervention[1] to be made
8		See few eligible patients
9		Lack of time
10		Involved in too many research projects
11		Insufficient financial compensation

[1] *The candidate understood these GPs to be referring to Shapiro's widely quoted definition of the placebo effect as any non-specific psychological or psychophysiological effect produced by a known inert therapy or one that is inappropriate for the specific condition of interest (Shapiro & Morris 1978).*

Table 46: Questionnaire items, response category construction and rationale for item inclusion *(pages 1 of 3)*

	Questionnaire items	Response category construction	Rationale for item inclusion
1	**Did any of the following factors reduce your willingness to participate in the Acupuncture for Menorrhagia trial?**	**"yes" "no" "possibly"** "circle one number on each line"	
A	Menorrhagia is not an area of specific interest for you		4;7;9;11
B	Acupuncture is not a therapy of interest to you		4;7;9;11
C	You see few eligible patients		5 ACUMEN
D	Patients had only a 50% chance of being allocated to the acupuncture group		4 YACBAC
E	Your personal time constraints within the consultation		1-4 YACBAC
F	Your uncertainty of the benefits of acupuncture for menorrhagia		1;6 YACBAC
G	Your uncertainty about the acupuncture service itself		1;6 YACBAC
H	Your uncertainty as to whether acupuncture would be acceptable to patients		1;3;4,6; ACUMEN
I	Other ongoing treatment modalities		1;3;4,6; ACUMEN YACBAC
J	Difficulty for patients attending acupuncture clinics (?travel)		ACUMEN YACBAC
K	Dislike the trial design (pragmatic RCT)		7 ACUMEN
L	Involvement in experimental studies disrupts the doctor-patient relationship		12
M	Insufficient financial compensation		4;7 ACUMEN
N	Involved in too many research projects		1;5 ACUMEN
O	Other factors (please specify)		
2	**Did any of the following factors increase your willingness to participate in the Acupuncture for Menorrhagia trial?**	**"yes" "no" "possibly"** "circle one number on each line"	
A	Your belief in the potential benefits of acupuncture		4;7;9;11 YACBAC
B	Patient interest		ACUMEN
C	Your wish to support this research project		1;3;4;7;9-11 ACUMEN YACBAC
D	The involvement of a GP advisor		1;2;9;10 ACUMEN
E	Positive feedback from patients who have received acupuncture		ACUMEN YACBAC
F	Acupuncture provided an additional treatment option		1 YACBAC
G	Positive personal experience of acupuncture		ACUMEN
H	The pragmatic trial design		4 ACUMEN
I	Financial compensation		1;3;6;9;12;14 ACUMEN
J	ACUMEN Newsletters		3;14
K	Other factors (please specify)		
3	**Did you find the trial protocol clear and straightforward?**	**"yes" "no" "fairly"** "Tick one"	2;5;6;13 YACBAC
4	**Did you find the entry criteria for patients into the study clear and straightforward?**	**"yes" "no" "fairly"** "Tick one"	2;6;5,13 YACBAC
5	**Roughly what proportion of eligible patients did you ask to consider taking part in the trial? 100%, 75%, 50%, 25% or 10%?**	"Tick one"	1;3,4;6;7;10
6	**What do you think were your reasons for not asking all your eligible patients to consider taking part in the trial?**	"Tick all that apply"	

A	Time constraints		1-4 ACUMEN
B	Not sure if patient would find acupuncture an acceptable treatment option (too unusual)		1-6; ACUMEN
C	Patient distress (e.g. life events)		12
D	Disruption of doctor-patient relationship		12
E	Forgot!		3;5 ACUMEN
F	Your confidence in efficacy of available medical treatments		3; ACUMEN
G	Patient co-morbidity		
H	Patient likely to find attending acupuncture clinics too difficult (?travel)		ACUMEN
I	Other (please specify)		
7	**This question asks you to think about different recruitment methods. ACUMEN asked you to invite eligible patients to consider taking part during a consultation. Would you have found it easier to identify patients on your practice list and then have written to invite them to contact the researcher directly (if we had provided the materials and funding)?**	**"yes" "no" "possibly"** "Tick one"	11 ACUMEN
8	**Is there anything else we could have done to encourage you to ask patients to consider taking part in the trial? Please comment...**		YACBAC
9	**Over the course of the study we have tried to maintain a high profile with you. Has this been effective?**	**"yes" "no"** "circle one number on each line"	
10	**Do you recall.....?**		
A	ACUMEN Logo		10 YACBAC
B	ACUMEN study colour		10 ACUMEN
C	Phone calls from the researcher		1;3;7;11;14
D	Patient acknowledgement letters		3;7 YACBAC
E	ACUMEN Newsletters		3;14 ACUMEN
F	ACUMEN Posters for the surgery waiting room		ACUMEN
G	ACUMEN Information packs		1;2;10
H	Other (please specify)		
11	**How TRUE or FALSE is each of the following statements for you?**	**"definitely true" "mostly true" "not sure" "mostly false" "definitely false"** "circle one number on each line"	
A	I would feel confident recruiting patients for a trial looking at acupuncture for low back pain		ACUMEN
B	I would feel confident recruiting patients for a trial looking at acupuncture for menorrhagia		ACUMEN
C	I would feel confident recruiting patients for a trial looking at acupuncture for depression		ACUMEN
D	I would feel confident recruiting patients for a trial looking at acupuncture for migraine		ACUMEN

Table 46 References

1. Ward E et al. *Conducting randomised trials in general practice: methodological and practical issues,* 49 British Journal of General Practice. 919-922 (1999)

2. Borgiel A et al. *Recruiting family physicians as participants in research,* 6 Family Practice. 168-172 (1989)

3. Van der Vint D et al. *Practical aspects of conducting a pragmatic randomised trial in primary care: patient recruitment and outcome assessment,* 50 British Journal of General Practice. 371-374 (2000)

4. Ross S et al. *Barriers to participation in randomised controlled trials: a systematic review,* 52 Journal of Clinical Epidemiology. 1143-1156 (1999)

5. Peto V et al. *Factors affecting general practitioners' recruitment of patients into a prospective study,* 10 Family Practice. 207-211 (1993)

6. Hunt C J et al. *Do doctors know best? Comments on a failed trial,* 174 MJA. 144-146 (2001)

7. Lovato L et al. *Recruitment for controlled clinical trials: literature summary and annotated bibliography,* 18 Controlled Clinical Trials. 328-357 (1997)

8. Silagy C et al. *Factors affecting the level of interest and activity in primary care research among general practitioners,* 6 Family Practice. 173-176 (1989)

9. Foy R et al. *Clinical trials in primary care: targeted payments for trials might help improve recruitment and quality,* 317 British Medical Journal. 1169(1998)

10. Murphy E et al. *Will you help me with my research? Gaining access to primary care settings and subjects,* 42 British Journal of General Practice. 162-165 (1992)

11. Bell-Syer S et al. *Recruiting patients to randomised trials in primary care: principles and case study,* 17 Family Practice. 187-191 (2000)

12. Pringle M et al. *Editorial: Randomised controlled trials in general practice,* 311 British Medical Journal. 1382-1383 (1995)

13. Tognoni G et al. *Randomised clinical trials in general practice: lessons from a failure,* 303 British Medical Journal. 961-971 (1991)

14. Jonker P et al. *Randomised clinical trials in general practice,* 304 British Medical Journal. 508(1992)

4.1.4 Questionnaire development: Specific questions to test the survey hypothesis

In order to test the hypothesis that the key barrier to active GP participation in the ACUMEN trial was GPs unfamiliarity with acupuncture as a treatment option for conditions other than those characterised by pain, a question was devised that asked ACUMEN GPs to rate how true or false they believed a statement to be about their confidence in recruiting patients to acupuncture trials for four different health conditions: low back pain, menorrhagia, depression and migraine [*4.2.1 Table 46*]. The hypothesis being that the majority of GPs would agree that they would feel most confident recruiting patients to a trial looking at acupuncture for low back pain or migraine, as opposed to menorrhagia or depression. Both back pain and migraine are conditions commonly treated with acupuncture in private and NHS settings, and there have also been successful and recent studies of each condition conducted in the UK (Thomas et al. 1999; Vickers et al. 2004). Depression was used as the other non-pain condition as it is a condition that is currently being studied by acupuncture researchers in the US (Schnyer et al. 2001), and at the time of the survey development initial proposals for similar comparative studies in the UK had been submitted for ethics approval and funding (Thorpe et al. 2002).

4.1.5 Questionnaire development: Assessing an alternative recruitment pathway

In order to assess the acceptability of an alternative recruitment pathway, a question was devised that asked ACUMEN GPs if they would have found identifying eligible patients on their practice list and writing to invite them to contact the researcher directly, easier than asking them to take part in the trial during a consultation. An open question to ascertain what else ACUMEN GPs thought could have been done to encourage them to recruit patients was also included *[4.2.1 Table 46]*.

4.2 *The ACUMEN GP Questionnaire Survey*

4.2.1 Questionnaire items and rationale for their inclusion

The items included in the questionnaire are listed in Table 45. Where appropriate the relevant reference justifying their inclusion is given. Questions developed in line with

the aims and objectives of this survey are indicated by the acronym ACUMEN, whilst questions informed by, or drawn from, the York Acupuncture for Back Pain survey of participating GPs are indicated by the acronym YACBAC *[Table 46].*

4.2.2 Response format

The researcher chose a three point scale response category construction for the majority of questions concerned with attitudes and behaviours, as these tend to lie on a continuum (Streiner & Norman 1995) *[see questions 1,2,3,4,7 in Table 46].* Here a dichotomous response format would have risked introducing error because of the limited choice of response levels, and as a consequence a loss of information and corresponding reliability. Indeed, the consensus is that the middle response category in such attitude/opinion questions does not necessarily represent a position of neutrality, and should therefore be included (McColl et al. 2001a), and that not to would risk spuriously forced or created "yes" "no" responses. Including the middle response category also reduces the risk of non-response on account of too few options. However, the question that asked ACUMEN GPs to indicate their reasons for not inviting all their eligible patients to take part in the trial provided a list of possible factors and asked that they "tick all that apply". This runs counter to the consensus opinion that respondents should be asked to indicate that a factor either "does apply" or "does not apply", as there is otherwise a danger that a non-response can be mistakenly interpreted as "does not apply".

A dichotomous "yes" "no" category construction was used for question 10 as it was considered to be "non-threatening", in that it was not questioning the GPs participatory behaviour in ACUMEN, but was instead asking the GP to judge how successful the research team had been at maintaining a high study profile.

The question concerning attitudes and beliefs about recruiting to acupuncture trials for different conditions scaled responses on five levels "definitely true" "mostly true" "not sure" "mostly false" "definitely false". This format has been found to give reliable results (McColl et al. 2001a; Streiner & Norman 1995), and has the added advantage of allowing a Chi square test for linear trend to be performed to test for a statistically significance difference in responses according to the condition. This test

is able to embrace the complexity of a five level response over four items and produce a 5 x 4 Table.

4.3 *Sampling frame*

In order to maximise the representativeness of the sample and generalisability of the results, questionnaires were sent to all GPs in the York and Selby Primary Care Trust (n=170).

4.4 Strategies employed to maximise response rates

GP response rates in postal questionnaire surveys average 54%, whilst those of non-physicians have a mean response rate of 68% (Asch et al. 1997; McColl et al. 2001b). This is significant as experts have suggested that a response rate of 70% is desirable (Borg & Gall 1983; Fowler 1984) and a response of 60% acceptable, if results are to be generalisable to all members of the sample frame. Given the experience of previous researchers in this field, a target response rate of 60% was set for this survey of GPs. To this end, measures to maximise responses, that had been recommended in recent reviews concerning the conduct and design of questionnaire surveys in health care research, were undertaken (McColl et al. 2001a; McColl et al. 2001b):

a. **Timing**: the survey occurred during the month of March, so avoiding the peak holiday periods of December, July and August.
b. **Pre-notification:** The questionnaire was preceded by the third ACUMEN newsletter, which was sent to all GPs via the internal mail system. The newsletter explained the purpose of the survey and promised respondents feedback in the form of a summary of the results *[see issue 3, Appendix 11]*.
c. **Cover letter:** A concise, traditionally formatted covering letter highlighted the relevance of the survey. It was signed by key members of the research advisory group (by scanning signatures onto the approved letter) and then printed on to headed paper [*see Appendix 51*].
d. **Confidentiality**: In both the cover letter and on the survey itself, respondents were assured confidentiality through the use of a study ID. This is usual practice and is understood to mean that "*...only the researcher is able to link*

the numbered questionnaire to a named individual, and that individual responses are not revealed to a third party without explicit permission from the respondent concerned" (McColl et al. 2001a).

e. **Front page of questionnaire:** The front page contained the survey title, identity of the organisations carrying out the research, stated confidentiality, and highlighted the survey "incentive" of a selection of six fine wines.

f. **Length of questionnaire and respondent burden**: The questionnaire design took into consideration the potential for response bias due to fatigue or carelessness in overly long questionnaires. Therefore GPs who participated in the study (ACUMEN GPs – that is those who had officially agreed to participate in the study, or had facilitated referral for individual patients who first contacted the researcher) were asked to complete a questionnaire that extended to just three sides of foolscap paper (n=43) [*see Appendix 52*]. GPs who had not participated in the ACUMEN trial (non-ACUMEN GPs), and were therefore likely to be less motivated to respond, were asked to complete a questionnaire that extended to just one side of foolscap paper and focused solely on factors that reduced their willingness to participate (n=127) [*see Appendix 53*]. The temptation to crowd questions onto a page or reduce the font size to reduce the apparent length of the questionnaire was avoided.

g. **Placement of questions within pages:** The questionnaire was organised so as to avoid splitting questions. Questions were ordered according to logic and a consistent design employed for similar question types. A vertical format was used throughout.

h. **Print details:** A distinct typeface with a font size of 12 points was used.

i. **Pagination**: To achieve a professional finish, the questionnaire for ACUMEN GPs was printed in booklet format (A3 folded). The questionnaire for Non-ACUMEN GPs involved double-sided printing onto a single sheet of A4 paper.

j. **Coloured paper**: To aid recall, the questionnaire was printed on study coloured paper.

k. **Prompt information**: To aid recall, a summary of the study with a flow diagram to illustrate the design was printed on the reverse of the questionnaire [*see Appendices 52 & 53*].

l. **Last page of questionnaire**: In the questionnaire for ACUMEN GPs, GPs were thanked for the time they had given to responding to the survey at the end of the questions. There was not space enough to thank respondents to the short version of the questionnaire.

m. **Reply envelope**: A stamped addressed envelope addressed to the research centre was enclosed. Postage was second class due to limited funds and a lack of evidence that first class post would improve the response rate,

n. **Incentive to participate**: Respondents were informed in the pre-notification newsletter, the cover letter and on the survey itself that they would be entered into a prize draw for a case of fine wines. The opportunity to win a case of champagne, as opposed to six opportunities to win one bottle, has been found to be a more effective incentive for GP surveys, with a difference in response rates of 15% (Thomson et al 2004).

o. **Follow-up contact (reminder)**: Non-responders were sent a second cover letter and survey questionnaire after two weeks.

p. **Feedback via ACUMEN Newsletter**: The final ACUMEN Newsletter informed GPs of the survey response rate and announced the GP prize draw winner.

4.5 Data management

The collection and management of data was carried out in line with the Data Protection Act 1998, and in accordance with the Policy for the Protection of Data and Data Providers prescribed by the Department of Health Sciences, University of York. Participants' contact details were entered into a master index at the research centre and password protected. To ensure confidentiality, a unique study number identified the questionnaires, and all data stored on computer and entered into SPSS for analysis was anonymised and coded with the unique study number.

5 Analysis

The total questionnaire response rate and response rates amongst ACUMEN as opposed to non-ACUMEN GPs were calculated and presented in Table form. The questionnaire data was analysed using the statistical package SPSS. The coded responses for questions one to seven and nine to eleven were counted and frequency

tables of the data produced. The function "split cases" was used to enable responses to the shared question (question one) to be displayed in separate tables and compared.

It was not unusual for respondents to give partial answers. The worst case involved twenty-one GPs from a large group practice returning their questionnaires in the same reply envelope after responding in the same way to the same one item in question 1 only. Consequently, a sensitivity analysis was conducted to assess the influence of these partial responses upon conclusions: The data for question one was first analysed treating these apparently partial answers as if complete answers *[Tables 48a to 50a]*, and then re-analysed treating a partial response as no response at all *[Tables 48b to 50b]*. Question 8 involved an "open question" and the one hand written response was recorded as a string variable and reported below. A histogram *[Figure 10]* was conducted to further demonstrate the statistical significance of the results for question 11.

6 Results

6.1 Response rates

One hundred and seven of the 170 GPs in the York and Selby PCT completed and returned the questionnaire within four weeks, giving a total response rate of 63%. The response rate was 65% (28/43) amongst ACUMEN GPs and 62% (79/127) amongst Non-ACUMEN GPs *[Table 47]*.

Table 47: Pattern of GP response rates to first and second mailings

	Response rate within 4 weeks of the Initial mailing	Number	Response rate at 1st mailing	Number	Response rate at 2nd mailing	Number
GPs	63%	107/170	49%	83/170	25%	22/87
ACUMEN GPs	65%	28/43	58%	25/43	17%	3/18
NON-ACUMEN GPs	62%	79/127	46%	58/127	27%	19/69

Analysis of GP responses to question one, treating partial answers as if complete answers, revealed that "*time constraints within the consultation*", seeing "*few eligible patients*", "*uncertainty of the benefits of acupuncture for menorrhagia*" and "*other ongoing treatment modalities*" were the four4 factors most frequently cited as having hindered participation, with more than half of the GPs indicating that these were definitely or possibly relevant factors *[Table 48a, items 1 to 4]*. "*Insufficient financial compensation*" was a factor that more than a third of respondents considered to have definitely or possibly affected their participation.

Table 48a: Factors that reduced GP willingness to participate in ACUMEN – *treating apparently partial answers as if complete answers*

	Factors in descending order of popularity with the most frequent response category per factor highlighted in bold	Yes Valid percent (number)	No Valid percent (number)	Poss Valid percent (number)	Total
1	Personal time constraints within the consultation	**61.4** (51/83)	26.5 (22/83)	12 (10/83)	83/107
2	See few eligible patients	**50.6** (41/81)	40.7 (33/81)	8.6 (7/81)	81/107
3	Uncertainty of the benefits of acupuncture for menorrhagia	**48.2** (40/83)	39.8 (33/83)	12 (10/83)	83/107
4	Other ongoing treatment modalities	**46.3** (37/80)	42.5 (34/80)	11.3 (9/80)	80/107
5	Insufficient financial compensation	30.7 (31/101)	**56.4** (57/101)	12.9 (13/101)	**101/107**
6	Uncertainty about the acupuncture service itself	29.1 (23/79)	**62** (49/79)	8.9 (7/79)	79/107
7	Uncertainty as to whether acupuncture would be acceptable to patients	23.8 (19/80)	**56.3** (45/80)	20 (16/80)	80/107
8	Difficulty for patients attending acupuncture clinics (?travel) *Responses for GPs approx. 15 miles from York (response rate 15/33)*	23.5 (19/81) *40% (6/15)*	**54.3** (44/81) *27% (4/15)*	22.2 (18/81) *33% (5/15)*	81/107 *15/15*
9	Menorrhagia is not an area of specific interest	21.5 (17/79)	**67.1** (53/79)	11.4 (9/79)	79/107
10	Involved in too many research projects	13.9 (11/79)	**75.9** (60/79)	10.1 (8/79)	79/107
11	Patients had only a 50% chance or being allocated to the acupuncture group	11.5 (9/78)	**76.9** (60/78)	11.5 (9/78)	78/107
12	Acupuncture is not a therapy of interest	7.5 (6/80)	**83.8** (67/80)	8.8 (7/80)	80/107
13	Involvement in experimental studies disrupts doctor-patient relationship	3.8 (3/79)	**84.8** (67/79)	11.4 (9/79)	79/107
14	Dislike the trial design (pragmatic RCT)	0 (0/73)	**94.5** (69/73)	5.5 (4/73)	73/107

Approximately a quarter of the GPs indicated that their participation had been hindered by "*uncertainties about the acupuncture service itself*" and its acceptability to patients, both in terms of a therapeutic option, and in terms of the practicalities of patients attending for treatment at the private clinics [*Table 48a, items 6 to 8*]. Fifteen of the 33 GPs practicing some 15 miles from York participated in the survey and all responded to the question that asked whether "*difficulty for patients attending acupuncture clinics (?travel)*" had reduced their willingness to participate in ACUMEN, and of these, 73% said it was a definite or possible barrier to participation.

Most GPs did not consider a lack of interest in either menorrhagia or acupuncture research as factors that had reduced their willingness to participate in ACUMEN *[Table 48a, items 9 and 12].* For almost all of the respondents, neither a "*dislike of the pragmatic trial design*", nor the fact that "*patients had only a 50% chance of being randomised to the acupuncture group*", were factors that had hindered their involvement *[Table 48a, items 14 and 11].* Their being "*involved in too many research projects*" already, or disliking the impact research involvement can have upon the doctor-patient relationship were also not factors that reduced the majority of GPs willingness to participate in ACUMEN, with over 75% responding negatively to these items *[Table 48a, items 10 and 13].*

6.3 Effect of partial answers upon results

When question 1 was reanalysed treating apparently partial answers as no answers at all, the sole difference in outcome was that the item "*insufficient financial compensation*" moved from 5th to 10th place in terms of response popularity, with almost 29% as opposed to around 44% reporting that it was or may have been a factor that reduced willingness to actively participate in the trial *[Table 48b].* Similarly, when the function split cases was used, this factor fell from 6th to 12th place amongst ACUMEN GPs, and from 5th to 9th place amongst Non-ACUMEN GPs. Because of the limited effect of partial responses upon findings, the corresponding Tables treating apparently partial answers as no answers at all are provided as Appendices *[Appendix 54, Tables 49b and 50b].*

Table 48b: Factors that reduced GP willingness to participate in ACUMEN – *treating apparently partial answers as no answers at all*

	Factors in descending order of popularity with the most frequent response category per factor highlighted in bold	Yes Valid percent (number)	No Valid percent (number)	Poss Valid percent (number)	Total
1	Personal time constraints within the consultation	**61.4** (51/83)	26.5 (22/83)	12 (10/83)	83/86
2	See few eligible patients	**50.6** (41/81)	40.7 (33/81)	8.6 (7/81)	81/86
3	Uncertainty of the benefits of acupuncture for menorrhagia	**48.2** (40/83)	39.8 (33/83)	12 (10/83)	83/86
4	Other ongoing treatment modalities	**46.3** (37/80)	**42.5** (34/80)	11.3 (9/80)	80/86
5	Uncertainty about the acupuncture service itself	29.1 (23/79)	**62** (49/79)	8.9 (7/79)	79/86
6	Uncertainty as to whether acupuncture would be acceptable to patients	23.8 (19/80)	**56.3** (45/80)	20 (16/80)	80/86
7	Difficulty for patients attending acupuncture clinics (?travel)	23.5 (19/81)	**54.3** (44/81)	**22.2** (18/81)	81/86
8	Menorrhagia is not an area of specific interest	21.5 (17/79)	**67.1** (53/79)	11.4 (9/79)	79/86
9	Involved in too many research projects	13.9 (11/79)	**75.9** (60/79)	10.1 (8/79)	79/86
10	Insufficient financial compensation	12.5 (10/80)	**71.3** (57/80)	16.3 (13/80)	**80/86**
11	Patients had only a 50% chance or being allocated to the acupuncture group	11.5 (9/78)	**76.9** (60/78)	11.5 9/78)	78/86
12	Acupuncture is not a therapy of interest	7.5 (6/80)	**83.8** (67/80)	8.8 (7/80)	80/86
13	Involvement in experimental studies disrupts doctor-patient relationship	3.8 (3/79)	**84.8** (67/79)	11.4 (9/79)	79/86
14	Dislike the trial design (pragmatic RCT)	0 (0/73)	**94.5** (69/73)	5.5 (4/73)	73/86

6.4 Factors that reduced ACUMEN GP willingness to participate in ACUMEN

Analysis of ACUMEN GP responses to question one, treating partial answers as if complete answers, revealed that more than half considered *"other ongoing treatment modalities"* and *"personal time constraints within the consultation"* to be the two key barriers to their active participation, and more than 40% cited their seeing *"few eligible patients"* as a definite or possible factor *[Table 49a, items 1 to 3]*. Over a third indicated that their participation had been hindered by their *"uncertainty of the benefits of acupuncture"* and their *"uncertainty as to whether acupuncture would be acceptable to patients" [Table 49a, items 4 and 7]*. A quarter of participating GPs indicated that patient referrals had definitely or possibly been hampered by their *"uncertainty about the acupuncture service itself"* and the potential practical difficulties their patients would face in attending for treatments *[Table 49a, items 5 and 11]*. Around one third gave *"insufficient financial compensation"* as a factor

PAGE NUMBERING AS ORIGINAL

6.5 *Factors that reduced Non-ACUMEN GP willingness to participate in ACUMEN*

For Non-ACUMEN GPs, the four key factors that definitely or possibly reduced their willingness to participate in the trial were *"personal time constraints within the consultation"*, their seeing *"few eligible patients"*, *"uncertainty of the benefits of acupuncture for menorrhagia"* and *"other ongoing treatment modalities" [Table 50a, items 1 to 4]. "Insufficient financial compensation"* was a definite or possible reason for non-participation for almost 50%, as were *"uncertainties about the acupuncture service itself"*, *"whether (it) would be acceptable to patients"* and concerns about the practical difficulties patients might face in attending for treatment *[Table 50a, items 5 to 8]*. Almost 40% cited, as a reason for non-participation in ACUMEN, the fact that *"menorrhagia is not an area of specific interest"*, although most (81%) were not uninterested in acupuncture research *[Table 50a, items 9 and 12]*. Being *"involved in too many research projects"* was a factor for around a quarter *[Table 50a, item 10]*. And again, neither concerns about disrupting the therapeutic relationship by participating in trials, nor a dislike of the pragmatic trial design, were factors that were considered to have hindered participation for the great majority (94%), although 25% of Non-ACUMEN GPs did indicate that the fact patients had only a 50% chance of being randomised to the option of acupuncture was a disincentive *[Table 50a, items 13, 14, and 11]*.

6.6 *Factors that increased GP willingness to participate in ACUMEN*

The data from question 2 of the survey revealed that the three key factors that led to GP participation in ACUMEN were first and foremost a wish to support the study (81%), second because *"acupuncture provided an additional treatment option"* (67%), and third because of *"patient interest"* (67%), with a further 5 to 10% indicating that these were possible positive influential factors *[Table 51, items 1 to 3]*. Almost 50% cited *"positive feedback from patients who have received acupuncture"* as a reason, and just over 40% gave their *"belief in the potential benefits of acupuncture"* as a factor that definitely or possibly increased their willingness to participate in ACUMEN. Around 25% indicated that the pragmatic trial design and the *"involvement of a GP advisor"* were factors that did, or may have, positively influenced their decision to participate in ACUMEN. And almost one fifth (19%)

Table 50a: Factors that reduced Non-ACUMEN GP willingness to participate in ACUMEN - *treating apparently partial answers as if complete answers*

	Factors in descending order of popularity with the most frequent response category per factor highlighted in bold	Yes Valid percent (number)	No Valid percent (number)	Poss Valid percent (number)	Total
1	Personal time constraints within the consultation	**67.2** (41/61)	19.7 (12/61)	13.1 (8/61)	61/79
2	See few eligible patients	**54.2** (32/59)	35.6 (21/59)	10.2 (6/59)	59/79
3	Uncertainty of the benefits of acupuncture for menorrhagia	**54.1** (33/61)	31.1 (19/61)	14.8 (9/61)	61/79
4	Other ongoing treatment modalities	**46.6** (27/58)	43.1 (25/58)	10.3 (6/58)	58/79
5	Insufficient financial compensation	35.1 (26/74)	**51.4 (38/74)**	13.5 (10/74)	**74/79**
6	Uncertainty about the acupuncture service itself	32.8 (19/58)	**56.9** (33/58)	10.3 (6/58)	58/79
7	Difficulty for patients attending acupuncture clinics (?travel)	28.8 (17/59)	**44.1** (26/59)	27.1 (16/59)	59/79
8	Uncertainty as to whether acupuncture would be acceptable to patients	25.9 (15/58)	**51.7** (30/58)	22.4 (13/58)	58/79
9	Menorrhagia is not an area of specific interest	24.6 (14/57)	**63.2** (36/79)	12.3 (7/57)	57/79
10	Involved in too many research projects	14 (8/57)	**75.4** (43/57)	10.5 (6/57)	57/79
11	Patients had only a 50% chance or being allocated to the acupuncture group	10.7 (6/56)	**75 (42/56)**	14.3 (8/56)	56/79
12	Acupuncture is not a therapy of interest	8.6 (5/58)	**81** (47/58)	10.3 (6/58)	58/79
13	Involvement in experimental studies disrupts doctor-patient relationship	5.3 (3/57)	**82.5** (47/57)	12.3 (7/57)	57/79
14	Dislike the trial design (pragmatic RCT)	0 (0/52)	**94.2 49/52)**	5.8 (3/52)	52/79

Table 51: Factors that increased GP willingness to participate in ACUMEN

	Factors in descending order of popularity with the most frequent response category per factor highlighted in bold	Yes Valid percent (number)	No Valid percent (number)	Poss Valid percent (number)	Total
1	Your wish to support this research project	**81** (17/21)	9.5 (2/21)	9.5 (2/21)	21/28
2	Acupuncture provided an additional treatment option	**66.7** (14/21)	23.8 (5/21)	9.5 (2/21)	21/28
3	Patient interest	**66.7** (14/21)	28.6 (6/21)	4.8 (1/21)	21/28
4	Positive feedback from patients who have received acupuncture	42.9 (9/21)	**52.4** (11/21)	4.8 (1/21)	21/28
5	Your belief in the potential benefits of acupuncture	38.1 (8/21)	**57.1** (12/22)	4.8 (1/21)	21/28
6	Positive personal experience of acupuncture	19 (4/21)	**81** (17/21)	0 (0/21)	21/28
7	The pragmatic trial design	15 (3/20)	**75** (15/20)	10 (2/20)	20/28
8	ACUMEN Newsletters	14.3 (3/21)	**85.7** (18/21)	0 (0/21)	21/28
9	The involvement of a GP advisor	9.5 (2/21)	**76.2** (16/21)	14.3 (3/21)	21/28
10	Financial compensation	0 (0/21)	**100** (21/21)	0 (0/21)	21/28

gave their *"positive personal experience of acupuncture"* as a reason for supporting the trial. Not one gave *"financial compensation"* as a factor that increased their willingness to take part in ACUMEN, and with regard to the ACUMEN newsletters that were sent to all GPs in the York and Selby PCT, 14% felt that they had positively influenced their decision to participate.

6.7 *The trial protocol*

When asked whether they found first the trial protocol and second the patient entry criteria to the study to be clear and straightforward, the majority answered yes (76% and 73% respectively), and a further 18 to 19% indicated that they found them fairly clear and straightforward *[Tables 52 and 53]*.

Table 52: GP views as to whether the ACUMEN protocol was clear and straightforward

Response in descending order of popularity	Valid percent	Number	Total
Yes	76.2	16/21	21/28
Fairly	19	4/21	
No	4.8	1/21	

Table 53: GP views as to whether patient entry criteria to the study was clear and straightforward

Response in descending order of popularity	Valid percent	Number	Total
Yes	72.7	16/22	22/28
Fairly	18.2	4/22	
No	9.1	2/22	

6.8 *Patient recruitment*

Only one GP stated that they had invited all of their eligible patients to consider taking part in ACUMEN, with approximately one third indicating they had invited 50% or more and two thirds indicating they had invited 25% or less *[Table 54]*. The two key reasons given were forgetfulness (68%) and *"time constraints"* (57%) *[Table 55, items 1 and 2]*. A further 18% cited their *"confidence in the efficacy of available medical treatments"* and *"patient co-morbidity"* as factors that militated against patient recruitment *[Table 55, items 3 and 4]*. Seven percent were concerned patients would have found attending for acupuncture treatment too difficult, 4% were unsure whether patients would have found acupuncture an acceptable treatment option, and a

further 4% were concerned an invitation would have disrupted the doctor-patient relationship *[Table 55, items 5 to 7]*.

Table 54: Approximate proportion of eligible patients GPs asked to consider trial participation

		Valid percent	Number	Total
100%		4.8	1/21	21/28
75%	***38.1% invited 50% or more***	14.3	3/21	
50%		19	4/21	
25%		33.3	7/21	
10%	***61.9% invited 25% or less***	4.8	1/21	
<10%		23.8	5/21	

Table 55: Reasons GPs did not ask <u>all</u> their eligible patients to consider trial participation

	Reasons in descending order of popularity	Percent	Number
1	Forgot!	67.9%	19/28
2	Time constraints	57.1%	16/28
3	Confidence in efficacy of available medical treatments	17.9%	5/28
4	Patient co-morbidity	17.9%	5/28
5	Patient likely to find attending acupuncture clinics too difficult (?time/travel)	7.1%	2/28
6	Not sure if patient would find acupuncture an acceptable treatment option	3.6%	1/28
7	Disruption of doctor-patient relationship	3.6%	1/28
8	Patient distress (e.g. life events)	0%	0/28

6.9 *Alternative strategies to facilitate patient recruitment*

Question seven of the questionnaire sought to gauge GP views as to whether or not identifying eligible patients on their practice list and writing to invite them to contact the researcher directly (materials and funding supplied) would have been easier than recruiting patients during a consultation. Responses were almost equally divided with 38% voting against, 33% voting for and 29% unsure whether or not they would have preferred this type of database recruitment *[Table 56]*. When asked if there was anything else the trial team could have done to encourage them to ask patients to consider taking part in the trial, one GP advised the use of *"a hooked reminder in the computer that is linked to the read code menorrhagia" [Table 57]*.

Table 56: GPs views as to whether or not practice database recruitment strategies would have been more practicable

Response in descending order of popularity	Valid Percent	Number	Total
No	38%	8/21	21/28
Yes	33%	7/21	
Possibly	29%	6/21	

GPs views as to whether or not identifying eligible patients on their practice list and writing to invite them to contact the researcher directly (materials and funding provided) would have been easier than recruiting patients during a consultation.

Table 57: GP suggestions to encourage their active participation

Is there anything else we could have done to encourage you to ask patients to consider taking part in the trial?
-code menorrhagia."

6.10 *Trial management*

Questions nine and 10 of the questionnaire were designed to gauge whether ACUMEN had been successful in maintaining a high study profile amongst GPs. The global opinion amongst 50% of GPs was that ACUMEN had been successful in this aim, whilst 41% believed ACUMEN had been fairly successful *[Table 58]*.

Table 58: GP views as to whether or not ACUMEN was effective in maintaining a high study profile

Response in descending order of popularity	Valid Percent	Number	Total
Yes	50%	11/22	22/28
Fairly	40.9%	9/22	
No	9.1%	2/22	

The strategies that proved most memorable were the ACUMEN information packs (94.7%), the study colour (85%) and newsletters (85%). Sixty three percent recalled the posters for display in surgery waiting rooms, and 50% remembered receiving a letter about the outcome for patients who had agreed to consider trial participation. Forty five percent remembered the ACUMEN logo and phone calls from the researcher *[Table 59]*.

Table 59: GP recall of ACUMEN research strategies to maintain a high study profile

	Do you recall.... *(Factors in descending order of popularity with the most frequent response category highlighted in bold)*	Yes Valid percent (number)	No Valid percent (number)	Total
1	**ACUMEN Information packs**	**94.7% (18/19)**	5.3% (1/19)	19/28
2	**ACUMEN Study colour**	**85% (17/20)**	15% (3/20)	20/28
3	**ACUMEN Newsletters**	**85% (17/20)**	15% (3/20)	20/28
4	**ACUMEN Posters for the surgery waiting room**	**63.2% (12/19)**	36.8% (7/19)	19/28
5	**Patient acknowledgement letters**	**50% (10/20)**	50% (10/20)	20/28
6	**ACUMEN Logo**	45% (9/20)	**55% (11/20)**	20/28
7	**Phone calls from researcher**	45% (9/20)	**55% (11/20)**	20/28

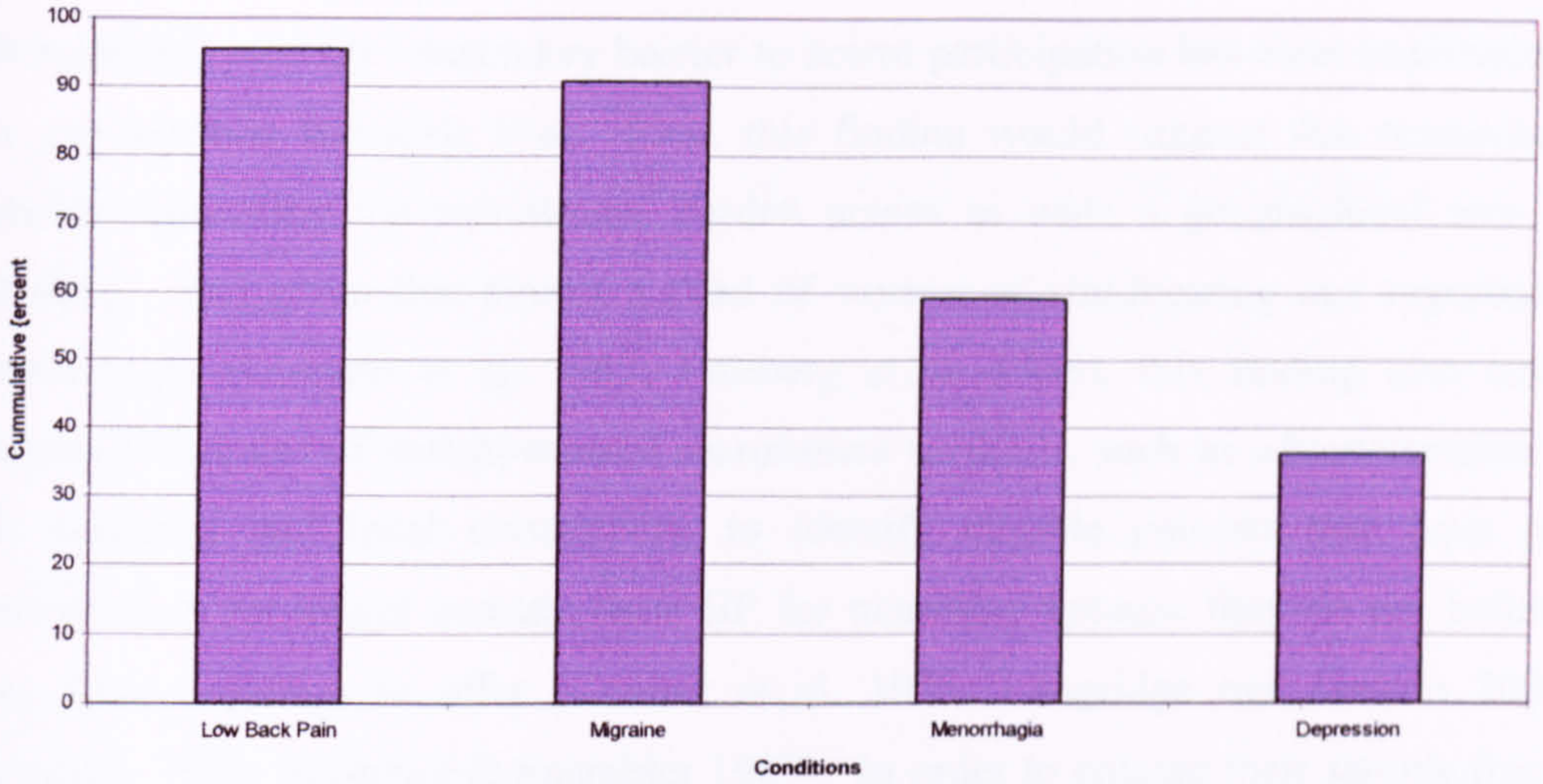
The data from the final question of the survey revealed that 68% of GPs would *"definitely"* feel confident recruiting patients for a trial of acupuncture for low back pain, with a further 27% indicating that they would feel *"mostly"* confident in this role *[Table 59, item low back pain].* Thirty six percent indicated that they would *"definitely"* and 55% that they would *"mostly"* feel confident if recruiting patients to a trial of acupuncture for migraine *[Table 60, item migraine].* When asked this question in relation to a trial of acupuncture for menorrhagia, 18% indicated that they would *"definitely"* and 41% that they would *"mostly"* feel confident if recruiting patients, whilst 18% indicated that for the most part, they would not feel at all confident *[Table 60, item menorrhagia].* In trials of acupuncture for depression, 14% of GPs anticipated that they would *"definitely"* and 23% that they would *"mostly"* feel confident if recruiting patients *[Table 60, item depression].* The majority (46%), however, were unsure and one responded that they would definitely not feel confident in this role. The cumulative percent for a positive response found that the great majority of GPs would *"definitely"* or *"mostly"* feel confident recruiting patients to trials of acupuncture for low back pain (96%) or migraine (91%), and more than half (59%) indicated that they would *"definitely"* or *"mostly"* feel confident recruiting to trials of menorrhagia. Less than 40% anticipated that they would feel confident recruiting for a trial of acupuncture for depression.

Table 60: GP anticipated confidence when recruiting patients to trials of acupuncture

I would feel confident recruiting patients for a trial of:	Definitely True Valid percent (number)	Mostly True Valid percent (number)	Not sure Valid percent (number)	Mostly False Valid percent (number)	Definitely False Valid percent (number)	Cumulative % for Mostly and Definitely TRUE (number)	Cumulative % for Mostly and Definitely FALSE (number)
Low Back Pain	**68.2%** (15/22)	27.3% (6/22)	0 (0/22)	3.6% (1/22)	0 (0/22)	**95.5%** (21/22)	3.6% (1/22)
Migraine	36.4% (8/22)	**54.5%** (12/22)	4.5 (1/22)	4.5 (1/22)	0 (0/22)	**90.9%** (20/22)	4.5% (1/22)
Menorrhagia	18.2% (4/22)	**40.9%** (9/28)	22.7 (5/22)	18.2 (4/22)	0 (0/22)	**59.1%** (13/28)	18.2% (4/22)
Depression	13.6% (3/22)	22.7% (5/22)	**45.5** (10/22)	13.6 (3/22)	4.5 (1/22)	**36.4%** (8/22)	18.1% (4/22)

Responses from 22 of a total of 28 GPs who returned their questionnaires
Conditions in descending order of popularity with the most frequent response category highlighted in bold.

Figure 10: Cummulative percentage of GPs who would definitely and mostly feel confident recruiting patients to trials of acupuncture for the treatment of four different conditions



7 Discussion

This survey benefited from acceptable response rates that exceeded the study target of 60%, both in total and within GP sub-groups. This suggests that the motivational strategies employed succeeded in enhancing responses above the average GP survey response rate of 54% (Asch et al. 1997), and that the results are therefore likely to be representative of GPs in the Selby and York PCT and may be extrapolated to other similar settings.

In part, the response rates may also reflect the finding that most GPs considered the York and Selby Acupuncture for Menorrhagia Exploratory Trial (ACUMEN) an interesting and valid study. Thus the results lend support to the pragmatic trial design itself, suggesting that GPs are interested in finding answers to practical questions about the extent to which an intervention produces a beneficial outcome under ordinary day-to-day circumstances – *"how well does it work in practice?"* - and are not only interested in studies that provide an answer to the questions *"is there a specific beneficial effect?"* or *"what is the active ingredient?"* (Thomas & Fitter 2002) - a finding that remains constant within GP subgroups.

However, the principle aim of this survey was to gain an understanding as to why ACUMEN had struggled to recruit patients. Thus, the fact that more than 40% of ACUMEN GPs and more than half of Non-ACUMEN GPs reported that their seeing *"few eligible patients"* was a key barrier to active participation has clear implications for any planned full-scale trial. First, this finding would suggest that researchers consider spreading the recruitment burden across as wide a geographical area as possible. And given that almost a third of women of childbearing age experience menorrhagia (Corrado et al. 1987; Hallberg et al. 1966), this finding also lends support to the use of patient-centred recruitment methods, such as advertisements in GP surgeries and local newspapers, to identify eligible patients that have not consulted, or no longer consult, their GP for treatment because they do not believe they have anything to offer (Coulter et al. 1995; Longridge nee Gamon 2000; Marshall 1998; Scambler & Scambler 1985). In order to enlarge their sample frame within the given population, researchers might also consider the additional use of retrospective recruitment methods, where eligible patients who consulted for treatment within the past 6 months, for example, are identified from the computerised practice list and sent a letter from their GP that invites them to contact the researcher directly. For, when asked whether they would have found this method easier than recruiting patients during a consultation, 33% of ACUMEN GPs responded positively and a further 29% indicated that they might have found this easier. The fact that 38% did not approve this strategy, however, would suggest it not be employed as the sole recruitment method.

Most of the 23 GPs who had initiated or facilitated patient referrals had done so for just one patient *[Table 61 below]*; and this was the case regardless of whether or not the GP had officially agreed to support the study. Six GPs had two to three patients enrolled and two GPs four and six. Given that the study had an 11-month recruitment period, these individual recruitment rates appear low, and it is relevant therefore that when asked about the proportion of eligible patients they had invited to consider trial participation, the majority (62%) reported their having invited less than 25%. However, eight of the 21 (38%) ACUMEN GPs responding to this survey question indicated that they had invited 50% or more, with one of these GPs asserting that they had invited all their patients and three that they had invited 75% of their patients. Thus this finding would seem to further emphasise the relevance of the barrier *"see*

few eligible patients". A knowledge of the number of patients with idiopathic menorrhagia registered as having consulted their GPs over the recruitment period gained through practice audit would greatly improve our understanding of how effective GPs were at recruiting and the relevance of this barrier to future research.

Table 61: Dispersion of trial patients amongst the GPs involved in their recruitment

GP	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	
Group*	A	A	A	A	A	A	A	A	A	A	A	A	
Pt	6	4	2	2	1	1	1	1	1	1	1	1	Total = 23
GP	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**		
Group*	B	B	B	B	B	B	B	B	B	B	B		
Pt	3	2	2	2	1	1	1	1	1	1	1		Total = 16

A=GP who formally agreed to support ACUMEN; B=GP who facilitated referral for 1 or more patients

"Uncertainty of the benefits of acupuncture for menorrhagia" was patently a key barrier to active participation amongst ACUMEN and Non-ACUMEN GPs, although this item received a greater number of positive or possible responses amongst Non-ACUMEN GPs. Indeed, almost 70% of Non-ACUMEN GPs believed that this was definitely or possibly an important block, as compared to the ACUMEN GPs who responded less emphatically to this item, with less than 40% believing that this was definitely or possibly a key reason for their reluctance to actively recruit patients to the trial. However, it is possible that ACUMEN GP responses to this question better reflect their wish to support the study and acupuncture research per se, rather than their beliefs about the potential benefits of acupuncture for menorrhagia. For an overwhelming 81% said that wishing to support the study was the key motivating factor for their involvement in ACUMEN, even though almost 60% said their involvement was definitely *not* motivated by a belief that acupuncture might be of benefit to patients with menorrhagia. They also cited "*other ongoing treatment modalities*" as the most important barrier to their active participation. Indeed, almost 60% indicated that this was a reason, or probable reason, for not actively recruiting more patients to the trial. And it may be, therefore, that *"other ongoing treatment modalities"* is another way of expressing *"uncertainty of the benefits of acupuncture for menorrhagia"*: Whereby a practitioner's lack of confidence in the experimental intervention is likely to lead to their favouring standard medical treatments. This is of import, as previous studies have found that a lack of confidence in the trial intervention, or greater confidence in an existing treatment, to be clear deterrents to

active GP participation in trials (Hunt et al. 2001; Ross et al. 1999; Van der Vint et al. 2000; Ward et al. 1999).

Thus the ACUMEN trial benefited from the goodwill of GPs, who recognised the need for this study of acupuncture and wished to demonstrate their approval by joining the trial. Yet ultimately the trial struggled to recruit patients because this GP support was severely tempered by uncertainty as to whether acupuncture would be of any real benefit to their patients with menorrhagia, which ultimately resulted in their preferring to continue to recommend standard treatment options and avoid introducing the study to patients. This uncertainty is a considerable barrier, and one that was undoubtedly compounded by a recruitment method that required GPs invite patients to participate in the trial during a consultation, for any discussion would risk exposing their uncertainties about the effectiveness of acupuncture for menorrhagia, and in turn risk undermining the integrity of the patient-practitioner relationship.

In addition, 30% of ACUMEN and 50% of Non-ACUMEN GPs indicated *"uncertainty as to whether acupuncture would be acceptable to patients"* to be a further disincentive for introducing the trial during a consultation. Thus, in the absence of GP perception that the trial provides patients with the opportunity to receive superior care, there is little incentive for them to generate referrals – unless of course the patient asks, when it becomes about facilitating a request and in so doing adding to the patients overall satisfaction with care, by providing them with a unique opportunity to receive a fully-funded course of acupuncture with an experienced professional acupuncturist at a private clinic. Under such circumstances, therefore, the experimental intervention does provide an attractive alternative to normal management, and it is worth noting that almost 70% of ACUMEN GPs cited the factors *"acupuncture provided an additional treatment option"* and *"patient interest"* as reasons for their involvement..

Thus these findings would seem to accurately reflect, and provide some explanation for why 35 out of 39 referrals to the ACUMEN trial were patient driven - primarily in response to surgery posters and a newspaper article about the study. They also provide an explanation for why not one pro-forma was faxed to the researcher to inform them of an eligible patient who had refused to consider trial participation. And

it would seem essential therefore, that any future trial acknowledges that GPs are often willing to refer willing patients to trials of acupuncture for menorrhagia, and that by adopting patient-centred recruitment strategies researchers can put in place recruitment pathways that are considerate of GP motives for supporting the trial.

However, it is unclear how future studies might achieve acceptable recruitment rates using patient-centred strategies alone, especially when we consider that many GPs declined displaying surgery posters, arguably due to the ambivalence described above. And therefore the principle challenge for any future study will be to transform GP interest and passive support into some kind of active participation; where the key barrier would seem to be GPs unfamiliarity with the use of acupuncture for menorrhagia, and perhaps more broadly, for conditions *not* characterised by pain.

The reason for asserting this perspective lies in the pattern of responses given to the question that asked ACUMEN GPs how confident they would feel recruiting patients to acupuncture trials for either low back pain, migraine, menorrhagia or depression. The great majority (95%) reported that they would definitely or mostly feel confident recruiting patients to a trial of acupuncture for low back pain, and a substantial 90% reported they would feel this way if it were a trial looking at the treatment of migraine headache, whereas if the conditions were menorrhagia or depression just 59% and 36% respectively would feel such confidence. Moreover, it is even possible to detect that GPs consider acupuncture best employed as a treatment for musculoskeletal pain as opposed to neurological pain. For, when asked to scale their responses on the following five levels: "definitely true" "mostly true" "not sure" "mostly false" "definitely false", the most popular response for low back pain was "definitely true" (68%), whilst for migraine it was "mostly true" (54%), with only 36% feeling fully confident of acupuncture's role here.

What was not anticipated, were GPs indicating that they would feel more confident recruiting patients to a trial of acupuncture for menorrhagia than they would to a trial of acupuncture for depression. The reason for not expecting this outcome being that it runs contrary to the popular conventional medical model which asserts that the effect of needle insertion is to stimulate peripheral nerves in muscles to send impulses to the central nervous system activating the release of endorphins, serotonin and

encephalins. These naturally occurring human biochemicals are compositionally similar to opiates. Thus they block pain perception, reduce inflammation and induce a feeling of well-being (Birch & Felt 1999; BMAS 2004; Pomeranz 2000). The biomedical community now broadly accepts this explanatory model, and it was therefore predicted that GPs would better conceive of a role for acupuncture in the treatment of depression than they would menorrhagia, because the effects of acupuncture upon the menstrual cycle are neither well known nor well understood, this being a relatively new field for western researchers (Stener-Victorin 2000). The reason for this unexpected survey finding may, therefore, be to do with the fact that this survey was conducted as part of an acupuncture for menorrhagia trial, which had the effect of biasing responses in favour of menorrhagia. A qualitative study involving semi-structured interview techniques would help us to better understand why GPs responded the way they did to this question.

However, it is worth noting that a previous study reported that its' recruitment success was partly due to the fact their GP information pack included a photocopy of a published feasibility study (Ward et al. 1999). And it is, therefore, entirely conceivable that this barrier will be less of an obstacle for future menorrhagia studies conducted in the UK and justified by the positive findings found in ACUMEN. For findings from this exploratory trial will provide GPs with a transparent rationale for active patient referral, regardless of their health beliefs, whereas prior to ACUMEN there existed no compelling evidence from effectiveness studies that acupuncture might be of benefit to patients. This new set of circumstances might also better allow GPs to express their uncertainties without calling into question their reasons for introducing the patient to the study, thereby maintaining the integrity of the patient-practitioner relationship.

Another factor found to militate against patient recruitment in previous studies was a lack of time, often because of the amount of paperwork GPs were asked to complete. Amongst ACUMEN GP's the factor *"personal time constraints within the consultation"* was second only to *"other ongoing treatment modalities"*, with almost 60% agreeing that this was definitely or possibly a reason for poor active recruitment rates. This is significant, as the ACUMEN protocol sought to minimise the time demands on GPs in the trial. GPs were required to first assess patients against the

inclusion-exclusion criteria, which primarily asked that they be reasonably confident that the woman's symptoms of menorrhagia were idiopathic in origin, second provide the patient with an information sheet to take home, and third gain the patient's written consent to fax their contact details, age and duration of symptoms to the researcher using the pro-forma provided. Moreover, the pragmatic trial design ensured that GPs would not be asked to comply with procedures that differed to normal practice, nor place any restrictions upon the treatment and care they normally provided their patients. And it is notable therefore that when asked, the majority of ACUMEN GPs indicated that they considered both the trial protocol and patient entry criteria to the trial to be clear and straightforward (76% and 73% respectively), with a further 20% considering the protocol and entry criteria fairly clear and straightforward. Therefore, it is possible that the number of positive responses to the barrier *"..time constraints.."* are less a reflection of the complexity of the trial and excessive amounts of paperwork, and more to do with the GP ambivalence described above. And it might be argued, therefore, that with the backing of a more sound evidence base, GPs would be better predisposed to give the time necessary to introduce patients to the trial within a consultation; this predisposition being more closely linked to GP confidence in the trial intervention than to a lack of time.

With regards to GP perception of the acupuncture provided by the trial, this survey revealed that *"uncertainty about the acupuncture service itself"* was a factor, or probable factor, that approximately 40% of non-participating GPs and 24% of participating GPs cited as having reduced their willingness to participate in the trial. The ACUMEN trial had sent all GPs in the Selby and York PCT a letter of invitation to participate and a study summary that informed them that *"patients randomised to acupuncture will be offered up to 20 treatments by one of three qualified acupuncturists with a minimum of three years experience working at a York-based clinical centre."* The clinic names and first line of address were then listed. No other written information was provided, although the training of the acupuncturists and locations of the clinics were often discussed during meetings with interested GPs. Thus this uncertainty may well have been due to the limited information with which GPs were provided, which would suggest that if they were equipped with detailed information about who would be providing the acupuncture and where, they would be more willing to refer patients to the study. Clearly, this is a factor that can be readily

addressed in future studies and might, for example, involve producing practitioner profiles comprising of a photograph, the qualifications awarded, training institutions attended and clinical experience to date. Information about the clinics might include pictures of the treatment rooms, a map to indicate their location and written instructions about how to reach them by car or bus. To ensure that this is a barrier attended to from the outset, researchers might consider including this information alongside the initial letter of invitation outlining the purpose of the trial and inviting GPs to participate.

Furthermore, by providing GPs with information about the practicalities involved for patients attending for acupuncture treatment, researchers might also address the concern about *"difficulty for patients attending acupuncture clinics (?travel)"*, which was a definite or possible factor that hindered trial participation for 56% of Non-ACUMEN GPs and 18% of ACUMEN GPs. Although this finding may also reflect the fact that patients in and around Selby and Easingwold were asked to travel some 15 miles to the practitioners providing the ACUMEN acupuncture service at clinics in York, and it is therefore worth noting that 73% of the GPs in these areas regarded *"difficulty for patients attending acupuncture clinics (?travel)"* a definite or possible barrier to their participation. Future studies may, therefore, need to consider how to provide as accessible an acupuncture service as possible for the patient population involved, perhaps setting up an acupuncture clinic within a participating GP practice for the purposes of the trial.

Future studies will also have to consider how best to address the issue of financial compensation. The funding arrangements for ACUMEN allowed for a small administration fee (£5.00) to be paid to GPs for each patient invited to consider participating in the trial. Predictably, not one GP gave *"financial compensation"* as a factor that had increased their willingness to participate. However, about one third of ACUMEN GPs and almost half of non-ACUMEN GPs gave *"insufficient financial compensation"* as a factor that had definitely, or possibly, hindered their participation. Indeed, whilst it was not unusual for respondents to give partial answers, the worst-case involved twenty-one GPs from a large group practice who returned their questionnaires in the same reply envelope after responding *"yes"* to this one item about funding. A stance that would suggest that their experience of *"insufficient*

financial compensation" had dominated all other considerations, to the extent that they were unwilling to engage fully in either the trial or related sub-studies such as this survey.

Overall, over 40% of GPs considered the token fee paid in ACUMEN to be a definite or possible barrier to participation. The percentage falls to just less than 30% when the responses from the 21 GPs in the large group practice are excluded from the analysis. Yet this is still a significant number, and there is no reason to conclude that the collective responses from GPs at the large group practice are in any way less representative of GPs in the PCT than those from questionnaires completed and returned independently. Thus, future studies will need to determine a fee that better reflects the time, thought and resources GPs will give to the study when facilitating patient referrals; whilst remaining mindful of the potential for scientific fraud, should the fee paid come too close to the threshold that demarks a *"financial inducement"* from a *"financial compensation"*.

In order to guard against poor participation, previous studies have highlighted the need to maintain a high study profile by employing strategies that enhance contact and communication between trialists and GPs (Bell-Syer & Klaber Moffett 2000; Foy et al. 1998; Jonker & Sumajow 1992; Lovato et al.1997; Murphy et al.1992; Ross et al. 1999; Silagy & Carson 1989; Van der Vint et al. 2000; Ward et al. 1999). This trial adopted purple as its' study colour and used the succinct 'ACUMEN' logo in a uniform way for letters, information packs, posters and newsletters. A small passport sized coloured photograph of the researcher on the top right hand corner of each newsletter was used to help familiarise GPs with the researcher and by extension the research itself. Almost all GPs remembered the information packs (95%) and the majority remembered both the study colour and newsletters (85%), which would suggest ACUMEN was successful in defining a memorable trial identity and keeping GPs up-to-date with trial progress; although conversely, only 45% remembered the logo. And whilst the deep purple plastic of the information pack made it memorable, a common observation made by GPs during meetings with the researcher was that because of its size, it was likely to become buried in a drawer or filed away on a shelf. Some GPs therefore commented upon the practical value of a desktop dispenser

containing A5 sized versions of the pro-forma and patient information sheet that might always be to hand, and so serve as a prompt.

The fact fewer respondents remembered the poster (63%) and patient acknowledgement letters (50%) would seem to accurately reflect first, that many GPs had declined to display the study poster, and second, that only 23 of the 43 ACUMEN GPs had received one or more patient acknowledgement letters. However, the fact that only 45% remembered phone calls from the researcher suggests that higher levels of personal communication may have been beneficial and could have reduced the proportion of eligible patients not invited to participate in the trial by their GP because of "forgetfulness". Indeed, 68% of ACUMEN GPs (19/28) gave *"forgot!"* as the principle reason for their failure to invite all their eligible patients to the trial. Yet it is unclear whether more persistent phone calls to encourage recruitment during busy clinic hours would have proven to be an incentive or disincentive, and the idea suggested by a survey respondent of a *"hooked reminder linked to the read code menorrhagia"* might, therefore, prove a more effective and less intrusive strategy to remind GPs of a study. However, this would only be possible in a large-scale trial with the funds to afford for EMIS to write a suitable program for their software.

In order to address the problem of forgetfulness, studies might also consider how best to establish effective channels of personal communication from the outset by carefully planning their GP recruitment procedures to ensure that the researcher meets all those GPs enrolled in the study. This is because it seems reasonable to assume that GPs would be more likely to welcome regular telephone calls from a researcher they have met and discussed the study with. For example, in some instances, an interested GP at a group practice would be responsible for meeting with the ACUMEN researcher and deciding whether or not the practice would participate. Thus, for those ACUMEN GPs not at the meeting, the researcher's follow-up phone calls would have been the first direct contact they'd received, and they may therefore, have been less disposed to accept the call, especially given the barriers to active GP participation, such as *"uncertain of the benefits of acupuncture for menorrhagia"*, discussed above. Therefore, under such circumstances, future researchers might be best advised to also ask to attend a practice meeting to give a short presentation, in order that they have an

opportunity to meet most of the GPs enrolled in the study and thereby improve the acceptability and effectiveness of follow-up phone calls, or even email.

Clearly, these researcher-led strategies to improve contact and communication could be easily implemented, whilst the *"hooked reminder"* would be a more expensive strategy with potential operational caveats and should therefore be considered carefully. Although, as GPs tend to see few women patients with idiopathic menorrhagia, their ability to establish an efficient, habitual, trial referral routine that would help counter problems with forgetfulness are less likely. Thus, with perhaps just 1 or 2 patients per month, a *"hooked reminder"* might be the only practical and effective strategy by which to improve the proportion of eligible patients invited to consider participation within a consultation.

Finally, although the majority of GPs approved the pragmatic trial design, approximately 20%, irrespective of subgroup, had definitely or possibly been put off by the fact *"patients had only a 50% chance of being allocated to the acupuncture group"*. This is a barrier that may be overcome in a large-scale trial where the power calculation dictates that a higher ratio of patients be randomised to the "option of acupuncture". Although this particular disincentive to active GP participation is unlikely to prove unduly problematic in a future trial where the majority of the other barriers identified in this survey have been successfully addressed.

7.1 Study weaknesses

This survey achieved an overall response rate of 63%, with 49% of GPs responding to the first mailing and 25% of initial non-responders returning the follow-up questionnaire. This is a good response for a survey of GPs and therefore hints at a missed opportunity, as a third mailing of the short questionnaire to all non-responders at 4 weeks would have been practicable. Minimal additional funds and time could have generated a response rate closer to 70% and so improved the generalisability of the results for the key question regarding factors that hindered GP participation in ACUMEN; it is of note that a survey of GPs to understand their attitudes towards the provision of acupuncture on the NHS achieved a response rate of 83% after 3 mailings (Lipman et al. 2003).

With regard to the generalisability of the results, it should be noted that GPs in the York and Selby PCT work in the locality of a reputable acupuncture training college and clinic, numerous private clinics, and have recently participated first in a large NHS funded trial of acupuncture for low back pain and second this university based exploratory study of acupuncture for menorrhagia. It is therefore probable that the findings from this survey are not representative of GPs less exposed to acupuncture and acupuncture research.

Question 5 of the survey questionnaire elucidated that whilst two-thirds of ACUMEN GPs had invited 25% or less of their eligible patients to consider participating in the trial, a third had invited 50% or more. This coupled with data showing that the most patients enrolled by any one ACUMEN GP was six and the usual number one, over a recruitment period of 11 months, would seem to stress the significance of the barrier *"few eligible patients"*, and yet the lack of any information about patients who may have declined their GP's invitation to consider participation cautions against concluding that eligible patients very rarely consult their GP. Thus poor question design compromised the comprehensiveness of the findings. A follow-on question asking what proportion of invited patients had accepted and declined, would have improved the usefulness of this data with regard to gauging the scarcity of eligible patients. As it stands, the question assumes GPs informed the research team about all those patients invited to participate in ACUMEN during a consultation, and does not allow for a situation where GPs chose not to conform to the study protocol that asked they fax a pro-forma to the research centre for all eligible patients invited to the trial, including those who had declined referral.

Poor question design compromised the reliability of responses to the question that asked ACUMEN GPs to indicate their reasons for not inviting all their eligible patients to take part in the trial. This question gave a list of possible factors and asked GPs to *"tick all that apply"*. Limiting GPs to just one response level may well have resulted in unreliable data and a loss of information about GP recruitment behaviour. The three-point scale response category construction that was used for the majority of survey questions concerned with attitudes and behaviours would have yielded more reliable and complete data.

7.2 *Recommendations for future research*

This survey of GPs identified five key barriers to their active participation in the York and Selby Acupuncture for Menorrhagia exploratory trial:

I. Seeing few eligible patients,
II. Unfamiliarity with acupuncture as a treatment for menorrhagia,
III. Uncertainty about the acupuncture service itself,
IV. Insufficient financial compensation, and
V. Forgetfulness.

In order to overcome these barriers and guard against insufficient recruitment rates in a full-scale trial, it is recommended that future research adopt a range of strategies: First, to attend to the problem of few eligible patients, a multi-centred trial is advised, as this would enable the recruitment burden to be spread across a larger geographical area and pool of eligible patients. Second, whilst the success of ACUMEN owed much to the strength of GP support and patient interest, ultimately it struggled to recruit patients because GP support was severely tempered by their unfamiliarity with acupuncture as a treatment for menorrhagia, and their subsequent confidence in and preference for existing treatments. To lessen the impact of this barrier it is advised that GPs be given a copy of the published exploratory trial. The rationale being that compelling evidence from a recent UK-based study will provide them with a transparent and logical reason for actively inviting patients to consider trial participation, and so greatly improve upon the GP initiated patient recruitment rates achieved in ACUMEN. Moreover, it is predicted that this evidence will allow GPs to express their uncertainties about the benefits of acupuncture for menorrhagia without either calling into question their reasons for introducing the patient to the study, or significantly deflating recruitment rates.

The barrier concerning GP uncertainty about the acupuncture service itself is one that could be easily remedied in a full-scale trial, requiring simply that GPs be provided with detailed information about both the practitioners and the clinics. Similarly, a more acceptable and realistic level of financial compensation could be discerned by finding out how similar NHS or MRC funded studies cost GP time and resources. And, given the issue of GPs seeing few eligible patients, an EMIS program to

establish a hooked reminder to the read-code "menorrhagia" is probably the most efficient solution to the problem of forgetfulness.

However, even with these strategies in place, the experience of running ACUMEN and findings from this survey would suggest that a full-scale pragmatic randomised trial of acupuncture for menorrhagia that relies solely upon GPs to initiate recruitment is not advisable. Rather, what is recommended is a multi-faceted recruitment strategy that is considerate of GP motives for supporting the trail, recognises patient interest and works to maximise the sample frame within the given population. Therefore, GPs would be asked to invite their eligible patients to the study during a consultation, but with the help of a hooked reminder to the read-code "menorrhagia", a desktop set of referral documents and eye-catching posters in the surgery and waiting room. To reach eligible women no longer consulting their GP for treatment, posters, newspaper advertisements and articles would ask women interested in participating in the trial to contact the researcher directly. And to reach women who have only recently consulted for treatment, or who are in between treatment reviews, GPs would be asked to write to eligible patients who consulted within the past six months and are listed on their practice database (materials, practical help from the researcher and funding provided).

7.2.1 Conclusions

Clearly the feasibility of any full-scale trial justified by the findings from ACUMEN, must be determined by the degree to which the barriers to active patient recruitment by GPs identified in this survey can be overcome and more efficient strategies developed. The findings from this survey provide the acupuncture research community with valuable evidence by which to develop recruitment protocols for full-scale pragmatic, randomised menorrhagia trials in primary care. However, whilst researchers can be fairly confident of their ability to factor in at the design stage successful strategies to overcome barriers such as GPs seeing few eligible patients, a lack of information about the acupuncture service, or insufficient financial compensation, what is less clear is just how effective the positive findings from ACUMEN will prove to be in turning passive GP support into active participation, both in terms of referrals during a consultation and the extent of GP willingness to

engage in retrospective database recruitment methods. It is therefore considered essential that some considerable thought and funding is given to the posters and advertisements, as it is possible that these patient-centred strategies could come to facilitate the primary recruitment pathway in a full-scale trial, just as they did in ACUMEN.

7.3 Proposed recruitment protocol for future research

This section details the proposed practical arrangements for engaging GP support and allocating patients to trial groups in a full-scale multi-centred pragmatic randomised trial to assess the clinical and economic benefits of offering acupuncture to patients with menorrhagia in primary care:

<u>Phase 1: Engaging GP Support</u>

All GPs from the Primary Care Trusts involved will be sent a concise, traditionally formatted letter highlighting the relevance of the study, stating ethics committee approval, and with reference to the published exploratory trial. It will be signed by key members of the research team (by scanning signatures onto the approved letter) and printed on headed paper. The letterhead will comprise the identity of the organisations carrying out the research and trial logo. This information will be presented in a uniform way on all documents produced for the study (study summary, information sheets, newsletters etc.) The trial will also adopt a study colour to further define a memorable trial identity. The following documents will be enclosed with the letter of invitation:

- ***Copy of the published exploratory trial***
- ***Study summary***: A concise study summary will be printed on A4 sized paper. On one side the title, aim, objectives and study design will be outlined. Particular reference will be paid to the inclusion/exclusion criteria, patient referral procedures and the financial compensation GPs will be awarded for each completed pro-forma faxed to the research centre, or letter of invitation sent to eligible patients identified on the computerised practice list. A flow diagram to

illustrate the trial design will be printed on the reverse and contact details for key members of the research team, including the GP advisor, will be given.

- ***Information leaflet about the acupuncture service***: Information about both the practitioners and clinics will be provided. Practitioner profiles will comprise of a photograph, qualifications awarded, training institutions attended and clinical experience to date. Information about the clinics will include pictures of the clinic exterior or treatment rooms, along with a map to indicate their location and instructions about how to reach them by car or bus.
- ***Reply postcard***: Those interested in participating in the trial will be asked to return a reply postcard, coded with the GP's identifier, stamped and addressed to the research centre.

The researcher will contact by telephone those GPs who returned the reply postcard to arrange a personal or practice meeting. The researcher will strive to meet all those GPs enrolled through meetings with a single GP representative for a group practice. The researcher will also telephone to invite those GPs who did not return the reply postcard to participate in the trial.

<u>Phase 2: Patient Recruitment</u>

Participating GPs will be provided with a trial information pack comprising of:

- A sturdy A5 sized desktop dispenser, in which to display the
- Patient information booklets and
- Copies of the patient referral pro-forma with the inclusion/exclusion criteria highlighted on the front cover.
- EMIS software installation CD for a hooked reminder
- Brief patient information sheet about the study,
- Reply slip and stamped envelope addressed to the researcher.
- One small and one large poster for display in the GP surgery and practice waiting room.

Patient recruitment pathway A

GPs will identify suitable patients during a consultation, with the help of a hooked reminder linked to the computer read code menorrhagia and posters displayed in the surgery and waiting room. Patients will be invited to consider participation in a trial that may involve the offer of acupuncture. GPs will send a pro-forma to the researcher with the patient's age and duration of symptoms of menorrhagia. GPs will obtain informed consent from patients to transfer their contact details (name, address and telephone number) to the researcher. Patients will be given an information booklet about the study. Where the patient declines to consider participating in the trial, the GP will record their reason.

Patient recruitment pathway B

Practice patients identified from the computerised practice list as having consulted for treatment for idiopathic menorrhagia within the past 6-months will be sent a letter from their GP inviting them to contact the researcher directly. Enclosed with the letter will be a brief information sheet, reply slip and stamped envelope addressed to the researcher. The researcher will send patients who respond a copy of the patient information booklet.

Patient recruitment pathway C

A weekly advertisement, similar to the study poster, will be placed in the healthcare section of local newspapers. Journalists from local and national newspapers will be invited to write articles about the study, and posters will be displayed in complementary health clinics, pharmacies and in ladies changing rooms at fitness centres. These will ask women with heavy periods to contact the researcher directly, if interested in participating in a trial that may involve the offer of acupuncture treatment. Those women who appear to meet the inclusion criteria will be asked to consult their GP. Where the patient's GP is not participating in the trial, the researcher will write enclosing the necessary documents for them to facilitate referral for an individual patient. A follow-up phone call will be made to the GP prior to the patient's planned consultation. The GP will send the pro-forma to the researcher and provide the patient with an information booklet about the study. Where the patient does not meet the entry criteria the GP will record the diagnosis. Should the patient decline to consider further participating in the trial, the GP will record the reason.

Patient recruitment pathways A, B and C

The researcher will contact all patients recruited to this stage of pathways A, B and C and obtain informed consent from those patients who decide to enter the trial. Baseline measures will be completed prior to randomisation. A remote randomisation service will ensure immediate and unbiased allocation to trial groups. The method of random permuted blocks will be use to ensure a relatively even flow of participants to treatment groups and conceal the random allocation code.

Phase 3: Maintaining GP support

The researcher will write to GPs to inform them of their patient's decision regarding trial participation, the outcome of randomisation, and the acupuncture practitioner consulted. At the end of treatment, the researcher on behalf of the acupuncture practitioner will send a letter describing the outcome from treatment. Regular study newsletters bearing a passport sized photograph of the researcher will be sent to all GPs in the Primary Care Trusts involved. The researcher will also make monthly follow-up telephone calls and send an email to participating GPs.

Chapter Six

Discussion: The ACUMEN programme of research

1. Findings from the ACUMEN programme of research

The descriptive systematic review presented in Chapter Two of the thesis revealed a lack of compelling evidence for acupuncture in the treatment of menorrhagia and the related idiopathic conditions primary dysmenorrhoea and irregular menstruation. Just two case-series exist that have sought to evaluate acupuncture in the treatment of excessive menstrual loss (Zhang 1994; Wangcheng 1988) and their positive findings are purely suggestive, as in addition to the bias and imprecision inherent in the study design, both suffer from a lack of methodological rigour whereby recruitment procedures and interventions are poorly defined and validated or credible outcomes lacking. The four studies that concerned the treatment of amenorrhoea or irregular menstruation suffered from similar problems (Ma 1999; Tureanu 1994; Mo 1993; Liu 1992), in spite of one being a controlled trial (Ma 1999). The majority of available studies address the treatment of primary dysmenorrhoea (12 of 18 studies), and most are case-series (10/12) (Ge 2000; Yu 1998; Jiao 1997; Wang 1996; Yong-hai 1995; Liang 1990; Zhan 1990; Wang 1987; Yuqin 1984; Steinberger 1981); with two pilot trials that indicate a possible beneficial effect from acupuncture (Helms 1987; Thomas 1985), although a lack of methodological rigour, such as inadequate randomisation procedures and inappropriate statistical tests, small sample sizes and evidence of confounding from carry-over effects diminish the potential usefulness of results. Thus this comprehensive review of the literature from across the evidence hierarchy revealed that well-conducted exploratory studies – from case-series to randomised trials – would usefully add to this evidence base, and that there is a real opportunity to make a substantive contribution to research in this area.

The York and Selby Acupuncture for Menorrhagia Exploratory Trial (ACUMEN) with nested qualitative investigation and complementary postal questionnaire survey was undertaken to explore the rationale for a definitive pragmatic trial to evaluate the relative clinical and cost effectiveness of different treatment options, including acupuncture within normal primary care. The trial is presented in Chapter Three of

the thesis, and despite the sample size of only 39, the findings are most promising, providing consistent, convincing evidence of genuine effects across measures of menstrual health and general health, with effects evident at 12-month follow-up. Of particular note is the significant difference between groups on the primary menorrhagia outcome (Aberdeen Menorrhagia Questionnaire) at six months that gave a very large effect size of more than 1.0; the statistically significant differences on the social functioning, pain and vitality dimensions of the SF-36 at 3-month follow-up; the especially large significant difference of 30 points on the pain dimension at six months; and the significant difference on the vitality dimension and mental component score of the SF-36 at 12-month follow-up. Moreover, these gains were reflected in patients reported satisfaction with treatment and treatment outcome and their stated treatment preferences; so reducing the possibility of an erroneous positive finding.

The nested longitudinal qualitative enquiry reported in Chapter Four of the thesis provides evidence of the acceptability of acupuncture treatment and how this evolved over time in line with lived experience for six of the women in ACUMEN. Baseline interviews revealed that what initially engendered treatment acceptability, and therefore credibility, was the perception of acupuncture as a safe and natural alternative to drugs and surgery, alongside positive experiences or reports of acupuncture *in other contexts*, as none had heard of acupuncture being used to treat menorrhagia: this is in direct contrast to GPs, whose unfamiliarity with acupuncture as a treatment for menorrhagia was a significant barrier to their active recruitment of patients to the trial. Once treatment began, this inquiry revealed that acceptability came to be determined by the complex interplay of components related to 'the acupuncture treatment *process'*, 'the holistic *perspective* of treatment' and 'the perceived *outcomes* from acupuncture':

I. Process: clinic environment, therapeutic relationship, sensation from the needles (i.e. pain and relaxation), time required (i.e. regular time for self versus difficulties juggling work and family commitments to accommodate treatment),
II. Perspective: treatment took into consideration full array of psychological and physiological signs and symptoms, learning to understand self as a whole, advice given to attain and maintain health,

III. Outcomes: normalisation of menstrual loss, improvements in other symptoms and well-being, positive changes in health behaviour.

Those aspects of *process* and *perspective,* in conjunction with, and more importantly mediated by *outcomes*, determined treatment acceptability and patients' willingness to adhere to the full-course of treatment. This interplay was expressed in terms of the opportunity cost involved in attending for treatment: patients considered that the gains (e.g. caring and supportive therapeutic relationship, learning to see self as a whole, and the relaxation and calm induced) outweighed the losses (time required and pain on needle insertion) where health benefits added significantly to quality of life. That acupuncture could be of benefit to concurrent symptoms and well-being, if not to the main complaint of menorrhagia, was considered a most important aspect of treatment acceptability and the rationale for treatment compliance where symptoms of heavy blood loss had persisted.

Consequently, the concept 'treatment acceptability' and the role these women had come to define for acupuncture in their future healthcare over the course of ACUMEN was not restricted to menorrhagia. Rather, what it was considered acceptable for – menorrhagia, certain other health problems, or both - was determined by the health gains experienced in ACUMEN. Because they could choose to continue to use acupuncture privately outside the trial, and because this option raised cost implications, treatment acceptability came to be expressed in terms of their willingness to pay out of pocket for acupuncture. And they were willing to pay for treatment where they could be relatively certain that the desired health outcomes could be achieved. Thus satisfaction with a course of fully-funded treatment as part of a trial can result in NHS patients buying acupuncture privately, in order to maintain the benefits experienced in the condition under investigation or to access what they perceive to be the most appropriate therapy for certain other health problems, based on lived experience. This investigation also found that, if acupuncture were ineffective for menorrhagia, the same pragmatism that led women to try acupuncture, led them to return to their GP for a new drug therapy, with surgery (hysterectomy) continuing to be the least favoured option.

Despite the small sample size and owing to the methodological rigour and degree of congruence with previous work (Walker et al 2004; Barrett et al 2003; Patterson & Britten 2003; Gould & MacPherson 2001; Luff & Thomas 2000; Longridge *nee* Gamon 2000; Patterson & Britten 1999; Cassidy 1998), these findings help to explain why patient-centred recruitment strategies proved the most successful in ACUMEN, why all 20 women randomised to acupuncture completed their course of treatment, and what was the complex package of 'active ingredients' that patients randomised to traditional acupuncture experienced in ACUMEN.

ACUMEN achieved 39 of a target of 40 participants by responding swiftly to problems with recruitment. Initially largely dependent upon the willingness of GPs to refer suitable patients, ACUMEN introduced two further recruitment pathways that were patient centred and continued to ask for GP support beyond the initial target of 10. By the end of the extended 11-month recruitment period a total of 43 GPs had contributed to patient recruitment and had demonstrated that they were willing to facilitate referral for willing patients; with 35 of the patients coming to participate in ACUMEN in response to patient-centred strategies, and just four coming on the invitation of their GP during a consultation. The subsequent recruitment rate was just 3.6 patients per month. These findings have serious implications for a full trial reliant upon GPs, as it is not unusual for trials to be compromised or even fail on account of poor recruitment rates.

The postal questionnaire survey of GPs in the Selby and York Primary Care Trust presented in Chapter Five of the thesis was undertaken to understand the barriers to active GP participation in ACUMEN and ascertain whether they might be overcome, or more efficient strategies developed for a full-scale trial. This survey identified five key barriers to active patient recruitment by GPs in ACUMEN: seeing few eligible patients; unfamiliarity with acupuncture as a treatment for conditions not characterised by pain; uncertainty about the acupuncture service itself; insufficient financial compensation; and forgetfulness. On the basis of these findings, this study can recommend procedures to overcome these barriers and guard against failure to complete a full-scale trial. A full trial should: be multi-centred to spread the research burden across a larger geographical area and pool of eligible patients; provide GPs with a copy of the published exploratory study, as such compelling evidence would

give GPs a transparent and logical reason for active patient referral; provide GPs with detailed information about acupuncture practitioners and clinics; pay GPs compensation similar to other NHS or MRC funded studies; install an electronic reminder triggered by the Read code "menorrhagia" to counter forgetfulness; and adopt a multi-faceted recruitment strategy where referrals are initiated by (1) GPs during a consultation with the help of posters and electronic reminders, (2) posters and advertisements asking women to contact the researcher directly, and (3) GP letters to eligible patients on the computerised practice list.

2. Strengths and weaknesses of the ACUMEN programme of research

The systematic review was conducted in accordance with NHS Centre for Reviews and Dissemination guidelines and employs many of the review methodologies advised to minimise bias including: an extensive search strategy; the use of checklists to assess methodological quality; by extension, a checklist to assess the quality of interventions in studies of acupuncture; a data extraction sheet to facilitate study selection; and two reviewers to assess study selection and to apply the quality criteria. All these features serve to improve the precision of the review. It is worth noting however that the checklists used to grade studies from across the research hierarchy - case-series to randomised trial - were developed or adapted for the purposes of this review and have not been validated; that is they have not undergone the programme of tests that would be required to be sure they are reliable in their assessments and achieve construct validity. A strength of these checklists is that they were all derived from published sources; a weakness is that the published sources differ in their wording of shared quality criteria and there is therefore a lack of convergence across the assessment instruments. These concerns will become more pressing as acupuncture study quality improves. Finally, the review did not search the grey literature.

The ACUMEN exploratory trial with nested qualitative investigation and complementary postal questionnaire survey uses the framework for the development and evaluation of complex interventions to improve health proposed by the Medical Research Council (2000). The trial itself is a rigorously designed and conducted pragmatic randomised controlled trial that meets all the relevant quality criteria set out

in the CONSORT statement except assessor blinding, which is very difficult given the constraints of a PhD in health sciences. The pragmatic aims of the trial are entirely consistent with the House of Lords (2000) recommendations in that they provide evidence of the justification for and feasibility of a full-scale trial to evaluate the clinical and economic benefits of offering acupuncture to patients with menorrhagia within normal NHS management. Most importantly, by aiming to evaluate the genuine effects of everyday clinical practice, the trial's pragmatic aims are congruent with the principles of traditional acupuncture, as they allow for the tailoring of the intervention to the individual and the development of a beneficial therapeutic relationship. Thus ACUMEN preserves the integrity of the therapy under investigation and meets the requirements of *'model validity'* as set out by Jonas and colleagues (2002) for the evaluation of complementary therapies after the design and implementation of ACUMEN.

More specifically, this exploratory study provides some evidence that the primary menorrhagia outcome measure used (Aberdeen Menorrhagia Questionnaire, AMQ) is a valid instrument for primary care research and that its psychometric validation, along with full validation of the two secondary condition-specific measures devised to complement the AMQ, should be considered. Furthermore, it provides evidence of intervention effects for the purposes of calculating the power of a larger trial; and the excellent response rates, with attrition below 20% at each follow-up, indicate that the strategies to minimise loss to follow-up were successful.

The four key weaknesses of the trial concern first, its small sample size of 39 and the possibility that ACUMEN failed to characterise acupuncture as consistently effective treatment (Type 2 error); second, the fact the trial struggled to recruit patients; third, the potential bias from the way the women came to participate in the trial; and fourth, the failure of the adverse events questionnaire to delineate between acceptable and unacceptable events associated with needle insertion.

The first of these weaknesses reflects that ACUMEN was an exploratory study and not designed to provide definitive answers to questions about effectiveness; whilst the second has implications for the feasibility of a full trial and led ACUMEN to test two additional patient-centred recruitment strategies and conduct a complementary survey

to identify barriers to active patient recruitment by GPs. However, these successful recruitment initiatives introduced a new problem – that of "resentment bias"; for 35 of the 39 women in ACUMEN came to participate via patient-centred recruitment strategies, that is they were self-initiated, and almost all (38/39) would have chosen to receive acupuncture if given the opportunity. This raises questions about the effect of their preferences upon outcomes; whilst the placebo effect of patient preference was maximised for those randomised to their preferred option of acupuncture within normal NHS management, for those randomised to their less preferred option of usual care from their GP alone, the placebo effect was in essence minimised, or at least greatly reduced. The fact that more than a third (14/39) had previously used acupuncture and six were recruited by their acupuncturist also suggests that the women in ACUMEN may differ in important ways to a usual GP population.

Finally, problems with the adverse events questionnaire meant that ACUMEN was unable to report accurately upon the number of adverse events experienced by patients which, whilst not life threatening or requiring medical attention, caused real distress and reduced treatment acceptability.

The nested qualitative investigation was designed and conducted with attention to many of the criteria suggested for increasing validity in qualitative research including: design, conduct and analysis in accordance with defined frameworks; candidate's role and status made explicit; purposive sampling strategies; interviews with prolonged engagement (lasting 45–90 minutes) and repeated contacts (baseline, and three and six months); development and refinement of thematic index whilst coding transcripts; assertions about relevance of themes and association between ideas continually tested; emergent index independently verified; development of case and thematic charts; and findings clearly linked to the data. The congruence of findings with other research was demonstrated (Walker et al 2004; Barrett et al 2003; Patterson & Britten 2003; Gould & MacPherson 2001; Luff & Thomas 2000; Longridge *nee* Gamon 2000; Patterson & Britten 1999; Cassidy 1998), and threats to transferability were examined. However, these strengths, including a self-awareness about biases and assumptions, cannot detract from the key weakness of this study, which is its small sample size. Samples of 20 or more are usual for applied policy research designed to

culminate in actionable outcomes; a sample size of six lessens the transferability of findings even where there is evidence of reliable methods.

The strengths of the ACUMEN survey included the development of questions with attention to current evidence about factors thought to help and hinder GP participation in trials, and the use of a three or five point scale response category construction for the majority of questions concerned with attitudes and behaviours, as these tend to lie on a continuum. The survey also achieved an overall response rate of 63%, which is good for a survey of GPs, and so hints at a missed opportunity, as a third mailing at four weeks requiring minimal additional funds and time could have generated a response rate closer to 70% and so improved the generalisability of the results for the key question regarding factors that hindered GP participation in ACUMEN. On the generalisability of results, it should be noted that GPs in the Selby and York PCT work in the locality of a reputable acupuncture training college and clinic, numerous private clinics, and have recently participated in both a large NHS funded trial of acupuncture for low back pain and this university-based exploratory study of acupuncture for menorrhagia. So the findings of this survey may not be representative of GPs less exposed to acupuncture and acupuncture research.

3. Strengths and weaknesses of the ACUMEN programme of research relative to the existing evidence base

A Cochrane review to determine the effectiveness of transcutaneous electrical nerve stimulation and acupuncture for primary dysmenorrhoea published in 2002 (Proctor et al) reported on the two controlled trials of acupuncture that exist, but excluded case-series and the bulk of acupuncture's evidence base. The briefing paper produced by the Acupuncture Research Resource Centre two years earlier and entitled *Gynaecology and Acupuncture: the evidence for effectiveness* (Longridge *nee* Gamon et al 2000) included case series and was therefore far more comprehensive and reflective of the professions current evidence base, but did not employ any of the review methodologies advised to minimise bias such as an extensive search strategy; the use of checklists to assess methodological quality; or two reviewers to assess study selection and apply the quality criteria. Thus the systematic review presented in

Chapter Two represents the first attempt to synthesise evidence for the effectiveness of acupuncture in the treatment of menorrhagia.

The ACUMEN exploratory trial presented in Chapter Three of the thesis is perhaps the first rigorously designed and conducted randomised trial of acupuncture for a menstrual disorder, and certainly for menorrhagia. Despite its small sample size of 39, as compared with the 215 participants in a study by Ma (1999), the ACUMEN trial is the first in this field to give compelling evidence of effect and justify a full trial.

The nested qualitative study presented in Chapter Four of the thesis is consequently the first longitudinal study of menorrhagia patients receiving traditional acupuncture within the confines of a trial, and whilst it demonstrates considerable congruence with prior research (Walker et al 2004; Barrett et al 2003; Patterson & Britten 2003; Gould & MacPherson 2001; Luff & Thomas 2000; Longridge *nee* Gamon 2000; Patterson & Britten 1999; Cassidy 1998), it also gives unique evidence about the way in which treatment acceptability changes over time for this patient population. This helps to explain the success of patient-centred recruitment strategies in ACUMEN and the 100% compliance with acupuncture. The findings will help the design and conduct of future trials in this field, especially the content of patient information sheets to enhance recruitment rates and the delivery of the complex acupuncture intervention.

There is evidence of the factors that help and hinder the recruitment of patients by GPs within a trial (Hunt et al. 2001; Bell-Syer & Klaber Moffett 2000; Van der Vint et al. 2000; Ross et al. 1999; Ward et al. 1999;; Foy et al. 1998;;; Lovato et al. 1997; Pringle & Churchill 1995; Peto et al. 1993; Murphy et al. 1992; Jonker & Sumajow 1992; Tognoni et al. 1991; Borgiel et al. 1989; Silagy & Carson 1989). A survey of GPs in a large NHS-funded trial of acupuncture for low back pain promises to provide insights into the factors that facilitated their successful recruitment of patients (Thorpe & Thomas 2002). Nevertheless the ACUMEN survey presented in Chapter Five is the first to provide evidence of the barriers of specific relevance to trials of acupuncture for menorrhagia. Moreover, these survey findings are likely to bear relevance to other trials of acupuncture for conditions *not* characterised by pain and for which there is a lack of credible evidence.

4. Interpretation of the findings in the ACUMEN programme of research in light of 2 and 3

In spite of the small sample sizes in the trial and nested qualitative study, by virtue of the rigour with which the entire programme of research was designed and conducted, ACUMEN achieves its goal of making a substantive contribution to the evidence base for acupuncture in the treatment of menstrual disorders and menorrhagia specifically.

5. Recommendations for future research

5.1 Future research planned by the candidate

Quantitative and qualitative assessment of outcomes used in ACUMEN

The two secondary measures devised to complement the Aberdeen Menorrhagia Questionnaire (AMQ) by focusing upon three symptoms of specific relevance to menorrhagia (York Menorrhagia-Relevant Profile – YMRP) and upon 15 common peri-menstrual symptoms (Peri-Menstrual Symptoms Questionnaire – PMSQ) demonstrated sensitivity to change and consistency with the AMQ and SF-36 dimensions of pain and vitality. Further research is planned by the candidate to test the measures' construct validity – that is the extent to which they are related to specified variables in accordance with an established theory or 'hypothetical construct' – for example the hypotheses that those with a severe severity score for pain and heavy bleeding on the YMRP and PMSQ have a severe severity score on the AMQ.

The way in which these measures usefully add to our understanding of menstrual health and treatment effects will also be considered further. Full validation of these measures in accordance with the steps prescribed by Streiner & Norman (2001) could be achieved to some extent within a full trial, although a specific study is the ideal. In the interim, the candidate plans to compile individual health outcomes profiles using the questionnaire and interview data collected at baseline, three and six months for six of the patients receiving acupuncture in ACUMEN. This analysis will enable the candidate to assess the extent to which these instruments encompassed and measured the full range of treatment effects patients reported at interview, with the aim of

informing the development of questionnaires for a full trial. Whilst outcome measures cannot capture all aspects of change, the evidence from previous research is that, if carefully chosen, they can register most of the important treatment effects most of the time (Patterson and Britten 2004; Long et al. 2000; Patterson & Britten 2000; Tully & Cantrill 2000; Hickey et al. 1996; Hill et al. 1996; Ruta et al. 1994). These findings will be prepared for publication in a peer-reviewed journal.

Assessing the rationale for a more tightly defined acupuncture protocol

ACUMEN focused on traditional Chinese medical (TCM) acupuncture as practiced by members of the British Acupuncture Council. The protocol required only that practitioners use acupuncture as the principle therapeutic intervention, and that they provide treatment according to the underlying syndrome patterns diagnosed –(i.e. according to the Zangfu). However, there were no restrictions on the use of any complementary diagnostic frameworks, such as Five Elements or Vital Substances, the principal request being that they provide a full and accurate record of treatment in the patient treatment booklets provided. Thus ACUMEN employed a fairly 'open' or 'pragmatic' protocol that allowed the practitioners freedom to use complex and individual approaches for different patients (MacPherson 2004).

A more detailed analysis of the treatments reported in ACUMEN is planned by the candidate in order to characterise: the similarities and differences in TCM diagnoses made within this medically defined population; the distribution of points used at a session; the key acu-points prescribed; and the frequency with which co-interventions and advice complemented the acupuncture. This analysis will consider whether a more precise protocol should be developed for a full trial – to facilitate replication without compromising generalisability or the integrity of the intervention. A more tightly defined TCM protocol was reported by Smith and colleagues (2002) for a study of acupuncture to treat nausea and vomiting in early pregnancy, and the authors of the York Acupuncture for Low Back Pain trial (YACBAC) are currently developing a flexible trial protocol with scope for individualised treatment (MacPherson et al. 2004). The findings from the qualitative investigation of treatment acceptability presented in Chapter Four will guide this process, together with an exploration of the skill base available to a full-scale trial of TCM acupuncture for menorrhagia.

Economic Protocol

The economic protocol for ACUMEN proposed to initiate the pilot of a primary incremental cost-effectiveness analysis with a secondary incremental cost-utility analysis (Drummond et al 1999). The study will take the perspective of patients and the NHS by counting and valuing those costs and consequences of providing treatment that are borne by either the individual woman or the NHS. The use and cost of private acupuncture during the trial will also be taken into account. Data were collected prospectively within ACUMEN to enable the candidate to value the full cost of each model of care. The main focus will be on the AMQ, a condition-specific, ordinal measure (Ruta et al 1995). The extra costs of providing referral to acupuncture and the additional health benefits gained will be compared in the form of a incremental cost-effectiveness ratio. The results will be expressed as the marginal cost per unit of effect. As there is evidence questioning the validity (Brazier et al 1993) and stressing the limitations (Drummond et al 1999) of the EQ-5D, the incremental cost-utility analysis will be undertaken using both the EQ-5D and the SF-36. The quality of life data from the SF-36 will be converted into utilities by mapping health states from the SF-36 onto the EQ-5D (Chancellor et al 1997). Results will be expressed in terms of cost per QALY, and compared with other interventions purchased by the NHS. Sensitivity analysis will explore the robustness of these results to plausible variations in key assumptions. Individual resource use profiles will estimate costs per patient, inform the sample-size calculation for the full trial, and thus ensure that adequate cost data are generated and the threat of skewness minimised. The completion and analysis of these data for publication in a peer-reviewed journal is among future research planned by the candidate. At the same time, the demographic data collected in ACUMEN will be compared with the CENSUS 2001 to assess the representativeness of the sample – important in generalising both effectiveness and cost-effectiveness.

5.2 Design considerations for exploratory trials of acupuncture

The survey conducted to identify barriers to active GP participation in the ACUMEN exploratory trial revealed that GP unfamiliarity with acupuncture in the treatment of menorrhagia is the key challenge to be faced in future trials. Moreover, this survey shows that this challenge would also threaten trials of acupuncture for depression, and

arguably any other condition *not* characterised by pain. This is therefore a finding of considerable import for researchers planning exploratory acupuncture trials in the absence of a compelling evidence base. Indeed, given the experience of ACUMEN where the recruitment period was extended from five to 11 months in order to recruit a target of just 40 patients, researchers might best be advised to consider alternatives to the gold standard of the RCT at this the exploratory stage. For example, a quasi-experimental, controlled before-and-after study could provide a viable and attractive alternative. Such a study could engage GPs at just two adjacent practices. New menorrhagia consultations, for example, would be monitored for 24-months. Over the first 12 months patients in both practices would receive standard care with measures taken at baseline, three, six and 12 months. In the second 12 months another group of newly consulting patients would follow the same protocol, except that those at one of the practices would have the option of acupuncture in addition to usual care from their GP. Thus, whilst potential bias is introduced, it is a design that minimises recruitment problems, and it might therefore be an acceptable compromise for some exploratory studies, as it may generate more readily the data needed to justify a full-scale RCT. Indeed, Wittmann and Walach (2002 p.95) argue that:

> *"...there is no such thing as a gold standard methodology. There is only a very cumbersome process of understanding a research context properly and finding the research methodology best suited to that context."*

However, ACUMEN does show that exploratory randomised trials of acupuncture for conditions not characterised by pain and that lack a credible evidence base are feasible, at least in York. The solution to recruitment problems was the use of patient-centred strategies such as surgery posters; newspaper features and advertisements; and GP letters to eligible patients listed on their practice database. A possible reason for success was that whilst GPs usually require condition-specific evidence of effectiveness, patients often consider acupuncture an attractive and credible treatment option on the basis of their perception of it as a safe and natural alternative to drugs and surgery, and alongside positive experiences or reports of acupuncture *in other contexts*.

5.3 Recommendations for a full-scale pragmatic randomised trial

5.3.1 Trial design and sample size

The ACUMEN trial results justify a full trial, and the complementary survey shows that this should be multi-centred to spread the recruitment burden across a large geographical area and pool of eligible patients. Calculation based on the estimated effect sizes from acupuncture suggests that, with attrition rates below 20% at each follow-up, a sample of 200 would suffice. But to be sure of detecting clinically significant changes on the AMQ over the longer term would need a sample of 400. Given experience in ACUMEN one might divide this target over 10 centres, with the aim of recruiting within 12 months.

However, it is possible that patient-centred strategies would facilitate the primary recruitment pathway in a full trial just as they did in ACUMEN, and so introduce "resentment bias" with the risk of exaggerating the acupuncture effect. A future trial might therefore, consider "waiting list controls", whereby those randomised to usual GP management alone in the first instance would receive a course of acupuncture after 12 months. However, with a 24-month trial, further loss to follow-up is inevitable, and requires a larger sample to maintain power. A further risk of patient-centred strategies involving newspaper articles and posters is that they can reduce the generalisability of findings to primary care populations. GP database recruitment methods to identify women who have consulted for treatment for menorrhagia might, therefore, be a more reliable strategy.

A definitive trial could test the effects of maintenance treatment for a subgroup of the patients randomised to acupuncture at baseline, in order to explore what might be the "optimum" package of care over the longer-term; maintenance treatments are normal within routine clinical practice and something that the patients interviewed were willing to buy. A study of this magnitude would also enable regression analysis to study whether TCM acupuncture diagnoses and chronicity are good predictors of treatment outcome, on the basis that no treatment is good for everybody (Witman & Walach 2002 p.102).

5.3.2 Strategies to enhance patient recruitment and maximise compliance with acupuncture

The findings from the survey culminated in a protocol for recruitment in a full-scale pragmatic randomised menorrhagia trial in primary care. Many of the strategies recommended are supported and enhanced by the findings from the qualitative investigation. For example, the qualitative study lends support to the use of patient-centred recruitment strategies in future trials, as it would appear these succeeded by tapping into a need for alternatives to the therapies currently available from GPs for menorrhagia, and an interest in, and preference for, acupuncture as a natural, low-tech option. Moreover, it is likely that future research will be able to enhance the effectiveness of methods such as GP surgery posters, letters from GPs and newspaper features, by reporting the findings from the ACUMEN trial and describing the treatment outcomes for the six women who participated in interview. However, as noted above *[5.3.1]*, the potential bias and impact upon generalisability that strategies such as posters and newspaper articles can introduce must be seriously considered.

The qualitative study would also suggest that potential barriers to participation, notably fears about needles and concerns about the amount of time that would be needed to attend for 20-sessions of acupuncture over six months, be addressed from the outset in the information sheet. For example, vignettes compiled from the data collected for this study would both illustrate, and provide a context for, information about needle sensations, making it more accessible and better able to allay fears. The positive and negative aspects of regularly giving time to treatment should be set out. And most importantly, the information sheet should describe traditional acupuncture's holistic approach and the range of benefits that patients often experience as a consequence, which can make attending for acupuncture worthwhile even where normalisation of menstrual loss is not experienced; again, vignettes or 'stories' compiled using the data from this study would be useful here.

Future trials should also look to minimise the time spent travelling to treatment by providing an acupuncture service accessible to the patient population. For some patients, participating in ACUMEN meant travelling 15 miles for treatment, and one of the three practitioners did not provide treatment after 5pm. However, the importance of the clinic environment and the location of acupuncture services in

future trials is less clear from this study, as only one patient referred to the effect of the treatment setting on their experience of acupuncture. Thus, researchers may wish to await the publication promised by Patterson and Britten (2004) about the effect of the institutional context of the intervention upon experience and outcome, before deciding whether the problem of access might be successfully addressed by providing acupuncture at both private clinics, as in ACUMEN, and in GP surgeries.

Future qualitative studies might usefully test the validity of the model *[Figure 8, Chapter Four]* developed to describe the way in which 'treatment acceptability' evolved over time in this study sample for both a larger sample of women receiving acupuncture for menorrhagia within a trial and for patients receiving acupuncture for other conditions. Researchers might also assess the relevance of the opportunity cost of time to treatment acceptability in other trials of acupuncture, and particularly in those that involve lengthy treatment courses. The implications of patients privately accessing maintenance treatment following a successful course of acupuncture should also be considered; especially to explore what might constitute optimum treatment in the long term, the real costs of the intervention, and where that burden falls.

5.3.3 Questionnaires – outcome assessment in a full trial

The questionnaires used in ACUMEN included the Aberdeen Menorrhagia Questionnaire and the short-form 36-item health questionnaire (SF-36). The Aberdeen Menorrhagia Questionnaire (AMQ) is a condition-specific measure that is responsive to small, but clinically significant changes in the patient's condition (Ruta et al 1995). The SF-36 is a popular health profile outcome measure that has been validated for use in the NHS and gives a distinctive profile for menorrhagia (Garratt et al 1993). Thus the AMQ and SF-36 have been found to be valid, reliable and responsive to changes in patients with menorrhagia. In addition, the ACUMEN questionnaires included: questions about demographic and gynaecological characteristics; the York Menorrhagia-Relevant Profile (YMRP), a measure that asked about three symptoms of specific relevance to menorrhagia – heavy bleeding, pain and fatigue; the Peri-Menstrual Symptom Questionnaire (PMSQ) about 15 common symptoms including breast tenderness and irritability, devised to estimate the wider effects of treatment upon menstrual health (PMSQ); two general questions that acted

as global indicators of change; questions about treatment priorities and satisfaction with care; questions devised to assess treatment acceptability and preferences; questions about side effects or responses to treatment; and questions about adverse events from acupuncture. Any future trial would be advised to re-word the questions about adverse events in line with those used in a recent survey of adverse events reported by patients (MacPherson et al 2004) in order to separate acceptable and unacceptable events associated with needle insertion. Future trials might also consider the addition of the Medication Change Questionnaire being developed by Patterson & Britten (2003) in recognition of the importance of medication avoidance or minimisation to women participating in ACUMEN.

5.3.4 Protocol for the comparator intervention

The active comparator intervention in ACUMEN was usual GP care; the aim was to assess the effectiveness of normal clinical practice and to maximise acceptability and participation. A risk of this strategy was that usual care would deviate too far from the guidelines for "best practice" and so compromised the study findings. In 1999, Prentice wrote *"despite widely available evidence inappropriate treatments are being prescribed"*, with only one in 20 prescribing tranexamic acid – *"probably the most effective first line treatment"* – compared with a third of GPs prescribing norethisterone – *"arguably the least effective option"* (IMS 1994). ACUMEN provides evidence that *"common practice"* still differs from *"best practice"* and suggests that the comparator treatment in a full trial need an explicit protocol.

6 Conclusion

A full-scale trial to evaluate the clinical benefits of offering traditional acupuncture to patients with menorrhagia in primary care is both justified and feasible. ACUMEN provides valuable information for the development of outcome questionnaires, effective patient recruitment strategies and patient information sheets, and for the delivery of the complex acupuncture intervention. However, questions about trial design remain: concerns about bias from patient-centred recruitment strategies; the rationale for "waiting list controls"; and more rigorous, but not rigid, protocols for the experimental and comparator interventions.