

**Acupuncture & moxibustion for
osteoarthritis of the knee:**
a component analysis approach

Ian Keith Appleyard

ORCID Identifier: 0000-0002-7012-3095

A thesis submitted in partial fulfilment of the
requirements of
London South Bank University
for the degree of Doctor of Philosophy

July 2018

Acknowledgements

I would like to thank my supervisors Professor Nicola Robinson, Professor Alison Twycross and Professor Nicola Crichton. Their support, encouragement and guidance throughout this project has been invaluable. I would also like to thank the many colleagues in the Confucius Institute and the School of Health and Social Care at LSBU who helped with the pilot study, offered advice on the practitioner survey, or answered my questions about Chinese language or specific points on Chinese medicine: Tian Ye, Zheng Yangkang, Dr Hu Xiaoyang (Mio), Martin Logue, Lynne Williams, Dr Huang Weidong, Yan Chao, Wang Zhijie, Zhang Weisong, Dr Liu Tingting, Yu Tianyang, Wang Liyuan, Chi Qingbin, Lyu Lusha, He Yibo, Song Dangyang, and Dr Esmé Paterson. I am grateful to Dr Stephen McKeever, Professor Alison Leary, Professor Martin Benwell and Dr Stephen Birch, for their advice.

I would also like to express my appreciation to Prof Xu Yinong, Executive Director of the Confucius Institute, for enabling the pilot study to be undertaken at the Confucius Institute clinic. Similarly, I would like to thank the British Acupuncture Council for providing a research grant for the pilot study. I would also like to acknowledge the participants in the pilot study, and those who selflessly gave up their time to be interviewed or complete the questionnaire.

And finally my thanks Huang Hailing for all her love and support, and Elizabeth Yueshan Appleyard for being a good listener.

Abstract

This project investigates the use of acupuncture and moxibustion for osteoarthritis of the knee. The project includes a pilot study which tests a protocol for a randomised controlled trial. Two pieces of primary research that develop new knowledge regarding acupuncture in practice: expert interviews and practitioner survey. A systematic review of clinical trials investigating warm needle acupuncture for osteoarthritis of the knee. A narrative review of the theoretical approaches to treating knee pain with acupuncture.

The first phase of the project reviews the literature on the methodological challenges of research into acupuncture. A new framework for acupuncture research is developed. Key issues identified are: a placebo acupuncture technique is inherently impossible; there is no 'theory of traditional East Asian medicine' that can be falsified. Component analysis is proposed as a potential solution to these challenges. Moreover, clinical research can only be well designed and properly interpreted if triangulated with knowledge of theory and practice.

The knowledge of acupuncture in practice, developed in this project, is used to re-evaluate the evidence-base. The external validity of acupuncture protocols and the risk of bias from the sham procedures are both shown to be problematic. This suggests that previous interpretations underestimate the efficacy of acupuncture. The research also indicates that acupuncture styles, e.g. TCM, cannot be clearly delineated, which has implications for evaluating and reporting clinical trials. In phase 3 of the project, the relationship between physical and psychological components is found to be highly complex. Subsequently, the research framework is further refined to account for this complexity. The pilot study identified problems with the protocol and solutions are proposed

The phases of the project are guided by the framework. Therefore, the project not only develops new knowledge regarding osteoarthritis of the knee, but also serves as a demonstration of the component analysis approach which could be utilised to investigate other conditions.

Table of contents

Acknowledgements	ii
Abstract	iii
List of boxes	xii
List of figures	xiv
List of tables	xvi
List of abbreviations	xix
Glossary of terms	xx
I. Introduction: a review of the challenges facing acupuncture research	1
I.1. Introduction	2
I.2. Paradoxes in acupuncture research	3
I.2.1. Paradox 1	4
I.2.1.a. Sham Acupuncture	5
I.2.1.b. Real acupuncture	6
I.2.2. Paradox 2	8
I.3. Traditional theories, knowledge and practice	9
I.3.1. Traditional East Asian Medicine (TEAM)	10
I.3.2. Historical diversity and authenticity	12
I.3.3. Inter-style homogeneity and intra-style heterogeneity	13
I.3.3.a. Defining styles used within research	15
I.3.3.b. Styles of practice in a state of flux	15
I.3.4. Moxibustion	16
I.3.5. The praxis of TEAM	16
I.3.5.a. Nosology - Bianbing	16
I.3.5.b. Diagnosis - bianzheng: a subjective process	17
I.3.5.c. Effectiveness may depend on individual skill	18
I.3.6. Summary	19
I.4. Placebo	20
I.4.1. Definitions of placebo and placebo effect	20
I.4.1.a. Definitions of placebo effect/ response	21
I.4.1.b. Definitions of placebo intervention	22
I.4.2. Interpretive bias and placebo	24
I.4.2.a. Risk of bias stems from the subjective definitions	24
I.4.2.b. Risk of bias from the false notion of a causal link	26
I.4.2.c. Risk of bias from conflating the spectrum of effects	27
I.4.2.d. Possible mechanisms ritual and expectation	28

I.4.2.e.	<i>The devaluation of psychological effects</i>	30
I.4.2.f.	<i>Ethical issues of placebo</i>	31
I.4.3.	Separation of psychological effects from acupuncture	32
I.4.4.	Are any acupuncture trials placebo-controlled?	34
I.4.5.	Should the term placebo continue to be used?	35
I.5.	Aims of the research	37
II.	Methodology: a novel framework for acupuncture research	38
II.1.	Introduction	39
II.2.	Epistemology	40
II.2.1.	The relationship between Science and TEAM	40
II.2.1.a.	<i>Epistemic relativism</i>	40
II.2.1.b.	<i>The demarcation of science and pseudo-science</i>	41
II.2.1.c.	<i>The ongoing re-evaluation of TEAM</i>	44
II.2.2.	Rationale for focussing on the physical components	46
II.3.	Acupuncture research frameworks	47
II.3.1.	Alternatives to the hierarchical methodology	47
II.3.2.	The components of acupuncture - schema	50
II.3.3.	Accommodation of TBSA clinical reasoning in research	51
II.4.	New research framework - Component analysis	55
II.4.1.	New schema of the components of acupuncture	55
II.4.2.	Micro - meso - macro theories in the new framework	59
II.4.3.	Key principle - independence of components	60
II.4.4.	Component efficacy - Risks and challenges	61
II.4.5.	Accommodation of clinical reasoning in the new framework	63
II.4.5.a.	<i>Optimal acupuncture</i>	64
II.4.5.b.	<i>Pattern differentiation - A staged process</i>	65
II.4.6.	Translational Research	70
II.4.6.a.	<i>Rationale for utilising TBSA</i>	70
II.4.6.b.	<i>Explanation of the need for a translational approach</i>	73
II.5.	The Project overview	76
II.5.1.	Phase 1: Acupuncture research (chapters-I/II)	76
II.5.2.	Phase 2 Literature review (chapters-III/IV)	77
II.5.3.	Phase 3 Acupuncture in practice (chapters-V/VI)	78
II.5.4.	Phase 4 Aims: Translational analysis and pilot study	79
III.	Narrative review: theoretical approaches	84
III.1.	Introduction	85

III.2.	Osteoarthritis of the knee	86
III.3.	Methods	87
III.3.1.	Selection of TCM textbooks	89
III.3.2.	Selection of non-TCM textbooks	90
III.3.3.	Methods of data extraction and analysis	91
III.4.	Basic acupuncture theory	91
III.4.1.	Bi-syndrome	92
III.4.2.	Pain 'mechanism' in TBSA	92
III.4.3.	Root and Branch - Ben and Biao	93
III.5.	Results - The standard theory	93
III.5.1.	Standard theory - The branch	94
III.5.2.	Standard theory - The root	96
III.5.3.	Replicability of the pattern differentiation	101
III.5.3.a.	<i>Replicability of the branch</i>	101
III.5.3.b.	<i>Replicability of the root</i>	102
III.5.3.c.	<i>Replicability - overall pattern differentiation</i>	104
III.5.4.	The standard theory - Ageing	105
III.5.5.	Replicability of treatment methods	105
III.5.5.a.	<i>Modes of acupuncture point selection</i>	106
III.5.5.b.	<i>Other physical variables</i>	108
III.5.6.	Summary - Standard theory	109
III.6.	Results - Alternative theories	111
III.6.1.	Alternative TCM Bi-syndrome theory	111
III.6.2.	Japanese acupuncture	111
III.6.3.	Medical Acupuncture	113
III.6.4.	Five Element - Stems and branches	114
III.7.	Discussion	115
III.7.1.	Limitations	115
III.7.2.	Assessment of osteoarthritis knee for the pilot study	116
III.7.3.	TEAM within a translational research approach	117
IV.	Systematic review: warm needle vs. needle acupuncture	119
IV.1.	Introduction	120
IV.2.	Evidence-base for needle acupuncture	121
IV.3.	Methods - Systematic review	124
IV.3.1.	Study Criteria	124
IV.3.2.	Search strategy	126

IV.3.3.	Data extraction and analysis	126
IV.3.4.	Risk of bias assessment	127
IV.3.5.	Quantitative data analysis	127
IV.4.	Results - systematic review	128
IV.4.1.	Included Studies	128
IV.4.2.	General characteristics	129
IV.4.3.	Risk of bias of included studies	131
IV.4.4.	Other Bias - Outcome measures	135
IV.4.4.a.	<i>Ordinal categories as outcome measures</i>	135
IV.4.4.b.	<i>Scale responder-rates outcome measures</i>	139
IV.4.4.c.	<i>Scale outcome measures</i>	143
IV.4.4.d.	<i>Risk of bias summary</i>	144
IV.4.5.	Participants	144
IV.5.	Analysis of outcomes	145
IV.6.	Searching Chinese language databases	146
IV.7.	Discussion	152
V.	Expert Interviews	154
V.1.	Introduction	155
V.2.	Methods	156
V.2.1.	The interview schedule design	156
V.2.2.	The interviewees and sampling	157
V.2.2.a.	<i>Selection criteria</i>	157
V.2.2.b.	<i>The recruitment process</i>	158
V.3.	Analysis	159
V.3.1.	Methodological approach - saliency analysis	159
V.3.2.	The framework method	159
V.4.	Ethics	162
V.5.	Results	163
V.5.1.	Distribution	163
V.5.1.a.	<i>Local acupuncture points</i>	164
V.5.1.b.	<i>Distal acupuncture points</i>	164
V.5.1.c.	<i>Number of points</i>	166
V.5.1.d.	<i>Adjustment</i>	167
V.5.1.e.	<i>Point location method</i>	167
V.5.2.	Stimulation	169
V.5.3.	Course of treatment	174
V.5.4.	Psychological factors	177

V.5.4.a.	<i>Anxiety</i>	177
V.5.4.b.	<i>Sensitivity</i>	177
V.5.4.c.	<i>Practitioner Belief</i>	179
V.5.4.d.	<i>Patient Belief</i>	179
V.5.5.	Finances and time	183
V.5.6.	Miscellaneous	183
V.6.	Discussion	185
V.6.1.	Replicability	185
V.6.2.	Emergent theme - psychological factors	187
V.6.3.	Limitations	188
VI.	Practitioner Survey	190
VI.1.	Introduction	191
VI.2.	Methods	192
VI.2.1.	Development of the questionnaire	192
VI.2.2.	Sampling frame	195
VI.2.3.	Distribution of the survey	196
VI.2.4.	Randomisation	197
VI.2.5.	Anticipated responses	198
VI.2.6.	Analysis	199
VI.3.	Ethics	199
VI.4.	Results	200
VI.4.1.	Response rates	200
VI.4.2.	Background information	202
VI.4.3.	Point distribution (section A)	204
VI.4.3.a.	<i>Number of needles - Questions 7-12</i>	205
VI.4.3.b.	<i>Modes of point selection - Question 13(a-i)</i>	208
VI.4.4.	Stimulation (section B)	216
VI.4.4.a.	<i>Needling variables (Q14-16)</i>	216
VI.4.4.b.	<i>Needle manipulation (Q17a-b)</i>	218
VI.4.4.c.	<i>Electro-acupuncture (Q17c-e)</i>	220
VI.4.4.d.	<i>Intra-dermal needles (Q17f)</i>	220
VI.4.5.	Stimulation - Moxibustion (section B)	221
VI.4.6.	Course of treatment (section C)	224
VI.4.7.	Adaption of treatment (section D)	227
VI.4.7.a.	<i>Psychological adjustment (Q24a-d)</i>	227
VI.4.7.b.	<i>Practical adjustment</i>	228
VI.4.8.	Miscellaneous - Auxiliary methods (section E)	230
VI.4.9.	Miscellaneous - Advice (section E)	231

VI.4.10. Safety	234
VI.4.1. Additional methods	235
VI.5. Discussion	235
VI.5.1. Summary	240
VII. Translational analysis: acupuncture for osteoarthritis of the knee	241
VII.1. Introduction	242
VII.2. Assessment Methods	243
VII.2.1. Previous assessment methods	243
VII.2.2. Establishing the criteria and benchmarks	244
VII.2.3. Methods of evaluating each criterion	247
VII.2.4. Selection of studies	248
VII.3. Re-evaluation of the evidence-base	249
VII.3.1. Assessment of external validity	249
VII.3.2. Assessment of the risk of bias from sham procedures	263
VII.3.3. Summary - external validity and risk of bias	268
VII.3.4. Implications for future research	269
VII.4. Generalisability of clinical research	269
VII.4.1. Generalisability assessed by modes of point selection	270
VII.4.2. Generalisability assessed by needling parameters	271
VII.4.3. Generalisability assessed by course of treatment	271
VII.4.4. Comparison of acupuncture points used	272
VII.4.5. Summary	273
VII.5. Comparison of theory and practice	279
VII.6. Discussion	280
VIII. Pilot Study: Randomised Controlled Trial	282
VIII.1. Introduction	283
VIII.2. Methods	286
VIII.2.1. Study Design	286
VIII.2.1.a. Randomisation	286
VIII.2.1.b. Blinding	288
VIII.2.2. Participants	294
VIII.2.2.a. Location of study	294
VIII.2.2.b. Eligibility Criteria	294
VIII.2.2.c. Recruitment	296
VIII.2.3. Interventions	298

VIII.2.3.a. <i>Intervention protocol - design process</i>	298
VIII.2.3.b. <i>Acupuncture protocol</i>	299
VIII.2.3.c. <i>Practitioners</i>	301
VIII.2.1. Data collection	301
VIII.2.1.a. <i>Quantitative outcome measures</i>	302
VIII.2.1.b. <i>Participant interviews</i>	302
VIII.2.1.c. <i>Patient expectation</i>	305
VIII.2.1.d. <i>The quality of blinding</i>	306
VIII.2.1.e. <i>Medication</i>	306
VIII.2.1.f. <i>Safety Evaluation</i>	306
VIII.2.1.g. <i>TCM pattern differentiation</i>	306
VIII.2.2. Analysis	306
VIII.2.3. Ethics	309
VIII.3. Results	312
VIII.3.1. Recruitment, attendance and data collection	312
VIII.3.2. Baseline characteristics	314
VIII.3.3. Participant Interviews	317
VIII.3.4. Researcher's reflection on the pilot study	324
VIII.3.5. Evaluation of blinding procedure	327
VIII.3.6. Expectation	330
VIII.3.7. Medication	332
VIII.3.8. Safety	334
VIII.3.9. Pattern differentiation	336
VIII.3.10. Outcomes measures WOMAC and RAND-36	336
VIII.3.10.a. <i>Change over time - repeated measures</i>	336
VIII.3.10.b. <i>Inter-group comparisons</i>	338
VIII.3.10.c. <i>Responder rates</i>	341
VIII.3.10.d. <i>Issues of concern with the outcome measures</i>	344
VIII.4. Discussion	344
VIII.4.1. Conclusion	348
IX. Discussion	349
IX.1. Introduction	350
IX.2. Phase 1: Methodological challenges and solutions of acupuncture research	351
IX.3. Phase 2: Literature based research	352
IX.4. Phase 3: Primary research - acupuncture in practice	355
IX.4.1. Overview of new knowledge from the practitioner survey	355
IX.4.2. Psychological factors influence the physical treatment	355

IX.4.3.	The symbolic nature of point selection	357
IX.4.4.	Physiological responses to the physical process may be moderated by psychological factors	358
IX.4.5.	Schema of acupuncture components - revised	360
IX.5.	Phase 4: Translational analysis and pilot study	362
IX.5.1.	Translational analysis	362
IX.5.2.	The pilot study	364
IX.5.3.	Final Summary	368
	References	369

List of boxes

I.1	Paradoxes in acupuncture research Langevin et al (2011)	4
I.2	Synopsis of Appleyard et al (2014)	7
I.3	Challenges from the characteristics of Chinese medicine	19
I.4	The Application of the Needle - Yellow Thearch's Internal Classic	32
II.1	Axioms of the new research framework	64
II.2	Attributes of bing-illnesses that maximise replicability	69
II.3	Summary of the attributes of the new research framework	75
II.4	Hypothesis of prospective RCT	80
II.5	Objectives of phases 2 and 3	82
II.6	Aims and objectives of phase 4	83
III.1	Objectives of the narrative review	86
III.2	Sub-patterns of bi-syndrome - Standard theory	95
III.3	Sub-patterns of bi-syndrome - Alternative TCM theory	111
III.4	Rationale for selection of osteoarthritis of the knee	116
IV.1	Objectives of the literature review	121
IV.2	Chinese language search strategy	126
IV.3	English language search terms	126
IV.4	Subject, abstract and title search terms	150
IV.5	Summary of phase 2 - criteria for prospective RCT	153
V.1	Objectives of the expert interviews	156
V.2	Additional psychological components of clinical reasoning	187
V.3	Key findings - of the expert interviews	189
VI.1	Objectives of the practitioner survey	192
VI.2	Adverse events - free text comments	234

VII.1	Objectives of the translational analysis	242
VIII.1	Aim and objectives of the pilot study	285
VIII.2	Inclusion criteria	294
VIII.3	Exclusion Criteria	295
VIII.4	Brief description of warm needle acupuncture	298
VIII.5	Expectation questions	305
IX.1	Summary of Phase 2: Literature based research	354
IX.2	Summary of Phase 3: Acupuncture in practice	361
IX.3	Summary of the translational analysis	363
IX.4	Summary of the pilot study	366

List of figures

I.1	Spectrum of factors encompassed by the placebo response	22
II.1	Sequence of research (based on Fønnebø et al., 2007)	48
II.2	Circular research methodology (Walach et al., 2006)	49
II.3	Acupuncture components China Academy schema (Luo et al 2015)	50
II.4	Acupuncture components SAR schema (Langevin et al., 2011)	51
II.5	Stages of the clinical reasoning process	53
II.6	Acupuncture components - new schema	56
II.7	Acupuncture components - expanded schema	58
II.8	Three hypothetical probability distributions for the effectiveness of individual acupuncture treatments	65
II.9	Pattern differentiation - Stages	68
II.10	Project overview	81
III.1	Illustration of root pattern differentiation	104
IV.1	Study selection flowchart	130
IV.2	Screenshot illustrating Chinese language search term issue	151
V.1	Stages of the treatment process - enhanced understanding	187
VI.1	Total number of needles used - on average	207
VI.2	Distribution graphs: A-shi, Local, Meridian and Relaxation	213
VI.3	Distribution graphs: Elbow, General health, Signs & Symptoms, Back shu	214
VI.4	Needle - depth, gauge, and retention time	217
VI.5	Stimulation graphs: Deqi, Tonify, EA-local, Intra-dermal	219
VI.6	Use of moxibustion distributions	223
VI.7	Course of treatment graphs	226
VI.8	Treatment modification graphs	229

VI.9	Auxiliary methods graphs	232
VI.10	Advice graphs	233
VII.1	Potentially active sham points	265
VII.2	Local points - citation frequency	274
VII.3	Distal points - citation frequency	275
VIII.1	Group allocation process - flowchart	287
VIII.2	Blinding procedure - flowchart	289
VIII.3	Close-up of warm needle acupuncture with skin guards	291
VIII.4	Treatment position with and without screen in place	292
VIII.5	Assistant lighting the moxa	293
VIII.6	Recruitment process - flowchart	297
VIII.7	Data collection process - flowchart	304
VIII.8	Recruitment, enrolment and data collection - flowchart	313
VIII.9	Stem and leaf plot: WOMAC sub-scale at baseline	315
VIII.10	Changes in the WOMAC Index over time	339
VIII.11	Changes in the RAND-36 PHC and MHC over time	340
VIII.12	Percentage changes in WOMAC global scores for individual participants	342
VIII.13	RAND-36 Individual MHC scores line graphs	343
IX.1	Schema of acupuncture components - revised version	360

List of tables

I.1	Definitions of the placebo effect/ response	21
I.2	Definitions of placebo interventions	23
III.1	Reviewed textbooks - TCM style	88
III.2	Reviewed textbooks - non-TCM styles	89
III.3	Standard theory - mechanisms and aetiology (abridged)	97
III.4	Standard theory - signs and symptoms related to the knee (abridged)	98
III.5	Standard theory - additional signs and symptoms (abridged)	99
III.6	Standard theory - treatment methods (abridged)	100
III.7	Comparison of knee osteoarthritis and bi-syndrome	102
IV.1	Short term pain outcome compared to 'sham' acupuncture	123
IV.2	Risk of bias summary - researchers judgement	133
IV.3	Category names	136
IV.4	Category outcome measures abridged	138
IV.5	Scale responder rates (abridged)	142
IV.6	Li Cd et al 2006 - Lequesne pain score change	145
IV.7	Zhang J 2013 - WOMAC pain score change	145
IV.8	VAS flat walking test pain	146
IV.9	Search field abbreviations CNKI specialist search	148
V.1	Interviewees - professional organisations and styles of practice	158
V.2	Framework Index	160
V.3	Distribution - synopsis	168
V.4	Stimulation - synopsis	172
V.5	Course of treatment - synopsis	176
V.6	Psychological factors - synopsis	181
V.7	Miscellaneous - synopsis	184
VI.1	Sampling frame - Professional Organisations	196

VI.2	Number of respondents for each allocated group	201
VI.3	Acupuncture association versus regulating body and style	203
VI.4	Allocated style of practice vs statutory regulating bodies	203
VI.5	Stem Questions - Point distribution Q13	204
VI.6	Number of needles used (Q7-12) Ranges and Kruskal-Wallis tests	206
VI.7	Number of needles used (Q7-12) Mann-Whitney U tests	208
VI.8	Q13 Stem questions - Kruskal-Wallis and medians	209
VI.9	Modes of point selection (Q13) Mann-Whitney tests	210
VI.10	Stem Questions - Stimulation Q17	216
VI.11	Stimulation (Q14-16) Mann-Whitney U tests	218
VI.12	Electro-acupuncture (Q17c-e) Mann-Whitney U tests	220
VI.13	Stem questions - Moxibustion Q20	221
VI.14	Moxibustion (Q20) Mann-Whitney U tests	224
VI.15	Stem questions - Course of treatment Q23	225
VI.16	Section C Course of treatment Mann Whitney U tests	227
VI.17	Stem questions - Adaption of treatment Q24	227
VI.18	Section D: Adaption of treatment Mann Whitney U tests	228
VI.19	Stem questions - Auxillary methods Q25	230
VI.20	Auxiliary methods Q25 Mann Whitney U tests	231
VI.21	Advice Q26 Mann Whitney U tests	234
VII.1	External validity assessment criteria - previous studies	244
VII.2	External validity assessment criteria	246
VII.3	Foster et al (2007) Non-insertion/ Semi-flexible	252
VII.4	Scharf et al (2006) Minimal/ Semi-flexible	253
VII.5	Witt et al (2005) Minimal/ Semi-flexible	254
VII.6	Williamson et al (2007) No sham comparison/ Semi-flexible	255
VII.7	Vas et al (2004) Non-insertion/ Fixed	257
VII.8	Berman et al (2004) Non-insertion/ Fixed	258
VII.9	Suarez-Almazor et al (2010) Minimal/ Fixed	259
VII.10	Tukmachi et al (2011) No sham comparison/ Fixed	260
VII.11	Hinman et al (2012) Sham laser/ Semi-flexible	261

VII.12	Mavrommatis et al (2012) Non insertion/ Fixed	262
VIII.1	Baseline demographic characteristics and pattern differentiation	314
VIII.2	WOMAC and RAND-36 (v1.0) scores at baseline	316
VIII.3	Summary of responses to questions regarding the recruitment process	318
VIII.4	Salient quotes from participants regarding the outcome measures	319
VIII.5	Participant interviews - acupuncture effects	322
VIII.6	Attitudes to acupuncture and perceptions of acupuncture	323
VIII.7	Participant guesses to group allocation questions and free text explanations	329
VIII.8	Expectations and mid-point pain - Fisher's exact test	331
VIII.9	Medication usage	333
VIII.10	Interviewee comments on adverse events	335
VIII.11	WOMAC and RAND-36 at the four time-points	337
VIII.12	Independent t-tests WOMAC global score	338

List of abbreviations

AACP	Acupuncture Association of Chartered Physiotherapists
ART	Acupuncture Randomised Trials [a specific series of trials in Germany]
ATCM	Association of Traditional Chinese Medicine
BAAB	British Acupuncture Accreditation Board
BACc	British Acupuncture Council
BMAS	British Medical Acupuncture Society
CAM	Complementary and alternative medicine
CER	Comparative Effectiveness Research
CIRC	Confucius Institute Research Collaboration
CONSORT	Consolidated Standards of Reporting Trials
CSA	Contemporary Styles of Acupuncture
GERAC	German Acupuncture Trials [a specific series of trials]
LSBU	London South Bank University
SAR	Society of Acupuncture Research
STRICTA	Standards for Reporting Interventions in Clinical Trials of Acupuncture
TBSA	Traditionally-Based Systems of Acupuncture
TEAM	Traditional East Asian Medicine
WMA	World Medical Association

Glossary of terms

Ben (root)	A component of the Chinese medicine diagnostic process that identifies an underlying cause for condition.	III.4.3
Bian bing	The process of identifying a key sign or symptom to categorise a patient's illness	I.3.5.a
Bian zheng	Can mean the diagnostic process of Chinese medicine or the diagnosis	I.3.5.b
Biao (branch)	A component of the Chinese medicine diagnostic process that identifies immediate cause of the pain in osteoarthritis (referred to a biao-branch in this thesis)	III.4.3
Bing	An illness in Chinese medicine (referred to a bing-illness in this thesis)	I.3.5.a
Chinese medicine	A cornucopia of concepts, explanations of physiology and pathology, therapeutic methods, and empirical observations originating from China	I.3.1
Complementary and Alternative Medicine (CAM)	A diverse group of therapies that are perceived to fall outside realm of conventional medicine	
deqi	A tingling or aching sensation typically experienced during acupuncture	
Effectiveness	A measure of the overall impact of an intervention	I.2
Efficacy	A measure of the impact of a specific component of an intervention	I.2
Hawthorne effect	The apparent improvement in a condition engendered by participating in a clinical trial. Part of the placebo effect	
Illness differentiation	English translation of bian bing	I.3.5.b
Minimal acupuncture	A type of sham or placebo acupuncture - normally needles are inserted less than 2mm in depth and at non-acupuncture point locations	
Needle acupuncture	Acupuncture using only needles, contrasted with warm needle acupuncture in this project	
Pattern differentiation	English translation of bian Zheng, referring to the process - also known as syndrome differentiation	I.3.5.b
Placebo	No clear definition	I.4.1

Placebo acupuncture	Procedures perceived to provide on the ritual aspects of acupuncture with no physiological action - also called sham acupuncture	
Practice	The use of acupuncture to treat patients, in non-research settings	
Praxis	Practice plus the theory and language used by acupuncturists to communicate with each other	
Project	All research and written work carried out as part of this PhD	
Qi gong	A mind body exercise that is part of tradition of Chinese medicine	
Sham acupuncture	Procedures perceived to provide on the ritual aspects of acupuncture with no physiological action - also called placebo acupuncture	
Syndrome differentiation	English translation of bian Zheng, referring to the process - also known as pattern differentiation	I.3.5.b
Tai ji quan	A mind body exercise that is part of tradition of Chinese medicine	
Type II error	Occurs when the null hypothesis is false, but erroneously fails to be rejected	
Warm needle acupuncture	A method of combining moxibustion with needle acupuncture	Box-VIII.4

I. Introduction: a review of the challenges
facing acupuncture research

导
论

I.1. Introduction

This project examines the use of acupuncture and moxibustion for osteoarthritis of the knee. There are three pieces of primary research; expert interviews, practitioner survey and pilot study. The expert interviews and the practitioner survey combine to gather information on acupuncture in practice. As far as the author was aware, at the time of writing, no similar studies have been published on osteoarthritis of the knee. Indeed, very little research has been conducted regarding acupuncture in clinical practice for any condition. The results and insights derived from all three pieces of primary research constitute new knowledge.

There is also literature-based research including: a narrative review of the theoretical approaches to the treatment of knee pain with acupuncture ([chapter-III](#)); a review the evidence base of acupuncture for osteoarthritis of the knee; a systematic review of warm needle acupuncture vs. needle acupuncture for osteoarthritis of the knee ([chapter-IV](#)). The systematic review was conducted using both English and Chinese language databases. The majority of the retrieved studies were published in Chinese, this is the first such review in English.

The project has two principal aims. To develop a protocol to investigate warm needle acupuncture, a specific method of acupuncture practice. To establish an overview of acupuncture for osteoarthritis of the knee that compared and contrasted three domains; theory, practice and clinical trials ([chapter-VII](#)). The overview enables a re-evaluation of the sham procedures and external validity of the acupuncture used in clinical trials. Thereby enhancing the previous understanding of the evidence base and identifying future avenues of research.

The literature on the methodological challenges in acupuncture research is reviewed. There is controversy regarding the interpretation of clinical research which has centred around the so-called paradoxes in acupuncture research

([section-I.2](#)). Two principal challenges are identified. The first challenge is a group of interrelated issues derived from the traditional theory and practice of acupuncture, these are examined in [section-I.3](#). The second challenge is the concept of placebo ([section-I.4](#)).

Whilst the methodological challenges are examined in this chapter, the proposed solutions are discussed in the next. A novel research framework has been developed as part of this project ([section-II.4](#)). The subsequent phases of this project are guided by the new research framework. Therefore, the aims and objectives of these phases are set out in [section-II.5](#), once the new research framework has been explained.

I.2. Paradoxes in acupuncture research

The evidence base for acupuncture has been typified by underpowered or poor quality clinical trials (Stux and Birch, 2001). Evidence sufficient to draw conclusions is generally limited to chronic musculoskeletal pain, with only a few exceptions such as post-operative nausea and vomiting (Cheong et al., 2013, Lee et al., 2015). A significant proportion of the high quality trials have been carried out in Germany - the Acupuncture Randomised Trials (ART), Acupuncture in Routine Care studies (ARC) and the German Acupuncture Trials (GERAC). The design of these studies has shaped the subsequent debates.

Different interpretations of efficacy and effectiveness lie at the heart of the debate. Effectiveness trials are pragmatic, comparing acupuncture to other interventions, usual care, standard care or waiting list controls. 'Effectiveness is a measure of the overall impact of an intervention outcome, as would be expected to occur in the routine care, with an emphasis on generalisability' (MacPherson and Hammerschlag, 2012, p143). The ethos of these trials is to understand whether the process of acupuncture *in its entirety*

will lead to clinically relevant benefits. Using this approach, it is not essential to understand whether any benefits are due to inserting needles or the psychological effect of the treatment. Efficacy has been defined as ‘a measure of the impact of an intervention on outcomes in as ideal conditions as possible, with the emphasis on controlling for placebo effects, thereby limiting the effects of bias’ (MacPherson and Hammerschlag, 2012, p143). In efficacy trials, it is of critical importance to understand what exactly caused any changes. In theory, by comparing acupuncture to placebo it should be possible to measure the proportion of clinical benefit that results from the needling, rather than psychological effects. The Society for Acupuncture Research (SAR) identified two key paradoxes (box-I.1).

Box-I.1: Paradoxes in acupuncture research Langevin et al (2011).

Paradox 1: A number of large well-designed clinical trials have reported that true acupuncture is superior to usual care, but does not significantly outperform sham acupuncture, a finding apparently at odds with traditional theories regarding acupuncture point specificity and needling technique.

Paradox 2: While many studies with animal and human experimental models have reported physiologic effects that vary as a function of needling parameters (e.g. needle insertion depth, mode of stimulation), the extent to which these parameters influence therapeutic outcomes in clinical trials is unclear.

I.2.1. Paradox 1

Paradox 1 stems from the apparent contradiction between effectiveness trials that favoured acupuncture and efficacy trials that had equivocal results. A simple explanation is: there is no difference between the real acupuncture and the sham acupuncture¹ because the superiority of acupuncture compared to usual care is due to non-specific factors. In essence, acupuncture is a placebo. Langevin et al (2011) proposed that an equally probable explanation was that a

¹ Sham and placebo acupuncture refer to the same procedures. However, it is important to consider the semantic implications of these terms, [section-I.4](#). Both terms are used throughout this thesis, as they are in the literature.

Type II error had occurred. They identified three plausible factors that needed further consideration.

- i. *Both real and sham/ placebo acupuncture have therapeutic effects*
- ii. *Acupuncture specific non-needle and/or needle components may be retained in sham / placebo treatments*
- iii. *That the real acupuncture provided within the context of acupuncture on trial is sub-optimal when compared to acupuncture in actual practice*

To understand the three factors, and how they have influenced the interpretation of results, it is necessary to elaborate on sham/placebo and real/verum acupuncture.

I.2.1.a. Sham Acupuncture

Fregni et al (2010) identified key challenges in Physical and Rehabilitation Medicine (PMR) clinical trials. At the top of the list was the development of the placebo/sham device. The development of placebos used in drug trials can be supported by knowledge of the biochemical processes and the physiological events likely to occur when a pharmaceutical is taken (Langevin et al., 2011). Ignorance of acupuncture mechanisms means there has been little understanding of what to avoid when designing placebo acupuncture procedures (Langevin et al., 2011, Wayne et al., 2009).

As part of the process of undertaking this PhD, an article *Should systematic reviews assess the risk of bias from sham–placebo acupuncture control procedures?* has been published (Appleyard et al., 2014) ([appendix-A.1](#)). It concluded that sham-placebo acupuncture techniques cannot be considered inert from either a biomedical or a traditionally-based systems of acupuncture (TBSA - [section-I.3.3](#)) perspective ([box-I.2](#)). It concurred with factor (ii) above that acupuncture specific needling components maybe retained in the sham controls. In other words, the effects seen in these sham-placebo treatments may be derived not just from psychological factors but also from the physical ‘sham’ needling process itself.

I.2.1.b. Real acupuncture

The question: 'What is acupuncture?' is more complex than it may at first seem. Acupuncture has not been clearly defined, a weakness in some systematic reviews stems from the heterogeneity of interventions classified as acupuncture. Needling at acupuncture points, needling at muscle motor points, trigger point needling as well as the use of lasers have all been classified as acupuncture (Ernst and White, 1998, White and Ernst, 1999). Trials investigating nausea have frequently used seasickness bands rather than acupuncture needles (Lee et al., 2015). Some interventions are relatively easy to exclude from the definition of acupuncture, such as lasers, others are more problematic.

Acupuncture could be defined, simply, as the insertion of needles at specific points on the body (Oxford Dictionary of English, 2005). Acupuncture could also be defined as a complex intervention that includes the consultation process and explanatory models based on traditional knowledge (Langevin et al., 2011). Luo et al (2015) emphasised the inclusion of the traditional diagnostic process that guides point selection; known as *bianzheng* pattern differentiation ([section-I.3.5.b](#)).

Clinical trials often simplify acupuncture treatment and may remove the traditional diagnostic process. Participants do not get the individualised treatment they would receive in usual practice. It is not known to what extent this affects clinical outcomes. The simplification suggests the acupuncture provided within the context of clinical trials maybe sub-optimal, when compared to acupuncture in actual practice (Birch et al., 2016, Langevin et al., 2011).

Box-I.2: Synopsis of Appleyard et al (2014)**Introduction**

Clinical guidelines depend on the analysis of randomised controlled trials in systematic reviews. How to interpret the results of sham–placebo procedures is a controversial aspect of the evidence. Two inferences can be drawn from the acupuncture vs. sham–placebo randomised controlled trials. The first is whether acupuncture has a physiological basis. The second is whether there is any validity in traditional concepts of acupuncture practice. The degree to which sham acupuncture controls can physiologically be considered placebo controls had already been challenged. The article evaluated the extent to which sham–placebo procedures should be considered inert, with particular reference to traditional acupuncture theories.

Results

The article identified four methods of sham-placebo acupuncture controls, superficial insertion, non-penetrating needles, use of non-acupuncture points, and points not indicated for a particular condition. These methods are used either individually or in combination.

Superficial insertion is part of traditional practice. Japanese acupuncture employs routine shallow needling of 1-2mm. In other styles depth of insertion may vary depending on the condition. Shallow insertion is used to treat muscle spasm caused by cold. Thus, in some cases deep insertion may be considered less effective.

Non-penetrating needles: It is not the case that penetration of the skin is thought to be essential. Again, in Japanese acupuncture non-penetrating techniques are used. Furthermore, there are massage techniques that use simple pressure on acupuncture points, such as tuina, shiatsu.

Non-acupuncture point: For painful conditions traditional theory recommends treating the muscle meridians. These follow the pathways of the main meridians but are wider. Sham points located a few centimetres away from a classical point may still be affecting the muscle meridian. Variation in point location between practitioners suggest sham points may overlap with true acupuncture points.

Points not indicated for a condition: In traditional practice it is rare to use a specific point in isolation for a particular condition. Therefore, it is unlikely that protocols using a single point will adequately reflect practice.

Conclusion

Sham acupuncture cannot be considered as an inert control from either a TBSA or biomedical perspective. There is a need to develop appropriate Acupuncture Control Assessment Guidelines to assess the risk of bias from sham–placebo controls when undertaking systematic reviews.

I.2.2. Paradox 2

The first explanation for this paradox is, simply, that basic science experiments have not provided information that is applicable to clinical research (Langevin et al., 2011). The majority of basic science research has investigated acute pain, induced in healthy adults or animal models. The physiological responses in chronic pain may be different to those observed in acute pain. Indeed, within TBSA acute pain maybe considered an *excess condition*, as opposed to chronic pain which may be a *deficient condition*². As a result, acupuncturists employ different methods and select different points to treat acute back pain when compared to chronic back pain (Legge and Vance, 1997, Maciocia, 1994). According to TBSA acute and chronic pain are not equivalent.

A second possibility is, in basic science research electro-acupuncture has been commonly used but in clinical trials manual stimulation has been frequently employed (Langevin et al., 2011). Some of the observed physiological responses were achieved only at high frequencies (100Hz) (Hammerschlag and Stux, 2001), which are not possible to mimic manually. A third explanation is there are no good biomarkers that can be used to measure any changes, as the pathophysiology of chronic pain is poorly understood (Langevin et al., 2011). Clinical research has been largely reliant on subjective outcome measures.

The issues of acute–chronic pain, electro–manual acupuncture, lack of objective biomarkers and limited understanding of the pathophysiology, have made it difficult to compare mechanism research with clinical trials. We cannot rule out the possibility that the needle-related physiological effects, seen in basic science experiments, are unrelated or only partially related to therapeutic effects.

Two factors related to TBSA also hinder the synthesis of mechanism and clinical research. First, there is lack of animal models that investigate the mechanisms of acupuncture within the explanatory framework of traditional theories. Secondly, it is unknown whether concepts such as acupuncture points,

² An explanation of the difference between ‘excess’ and ‘deficient’ can be found in most Chinese medicine introductory textbooks. The key point here is that they are different.

meridians and Qi have physiological equivalents that can be used to further develop understanding of acupuncture (Langevin et al., 2011).

Another explanation for paradox 2 ([box-I.1](#)) is that acupuncture needling does have therapeutic effects but these effects are relatively small when compared to the non-specific effects in clinical trials (Langevin et al., 2011). In order to reflect clinical practice, acupuncture within clinical trials is often more complex than that delivered in mechanism research. This heterogeneity may have obscured the small effects of needling.

I.3. Traditional theories, knowledge and practice

Acupuncture research not only examines whether or not there are physiological changes and clinical benefits, but also tries to evaluate the validity of the traditional theory. Some commentators have asserted that the evidence base demonstrates traditional theories are invalid (Moffet, 2009, Singh and Ernst, 2008). Acupuncture originates from a philosophical tradition different to that associated with science, it is in part this provenance that makes it a controversial intervention (Vickers et al., 2012). This narrative, which views acupuncture as either pseudo-scientific or a paradigm incommensurable with science, is discussed in [chapter-II](#).

One of the problems that has beset acupuncture research has been a lack of knowledge amongst the scientific community regarding traditional theories. This led to the design of trials that claim to be testing TBSA, yet, on examination, the treatments cannot be considered adequate in terms of traditional practice (Birch et al., 2004). If it is to be properly investigated, the traditional practice of acupuncture needs be understood. However, the challenges of incorporating TBSA into a modern scientific framework has not been well articulated within the current academic literature (Birch and Alraek,

2014). The objective of this section is to establish the barriers to incorporating traditional practice into a scientific framework.

I.3.1. Traditional East Asian Medicine (TEAM)

The first step towards incorporating the traditional practice and theories of acupuncture into a scientific framework is to develop an understanding the breath and nature of this knowledge. Birch and Felt (1999) put forward three general categories to summarise the complex of Chinese medicine: an uncapitalised 'traditional Chinese medicine', traditional East Asian medicine (TEAM) and a capitalised Traditional Chinese Medicine.

The first category, traditional Chinese medicine is the 'heterogeneous medicine of China's long recorded history' (Birch and Felt, 1999, p xiii). It encompasses the combined medical writings of thousands of authors and many different schools of thought. The foundational texts were created in the Han dynasty (Unschuld, 2009, Unschuld, 2010) - most notably The Yellow Thearch's Internal Classic³. These ideas have been transmitted and reinterpreted for two millennia. The concept of medicine can be referred to as 'a specialised form of healing because of its accessibility through written literature and privileged orthodox status' (Hinrichs and Barnes, 2013, p1). However, the medical and non-medical forms of healing, in China, have historically had a fluid interrelationship, and the two should be considered together in order to understand their development (Hinrichs and Barnes, 2013). So knowledge transmitted orally should also be considered part of traditional Chinese medicine.

The second category, TEAM, encompasses a broader understanding than 'traditional Chinese medicine' that includes other Asian countries. Medical texts and practices were transmitted to Japan in the sixth century A.D. (Veith, 2016). Subsequently practice evolved outside of China, and some styles of

³ 黄帝内经 *huangdi neijing* 'Yellow Emperor' is the conventional translation, however, 'Thearch' will be used in this thesis as it is a better translation of *huangdi* (Hinrichs and Barnes, 2013)

acupuncture and herbal⁴ medicine are now associated with Asian countries, such as Japan and Korea. However, delineating these styles is far from straightforward as: the praxis of Japan and Korean have their origins in the texts from ancient China (Veith, 2016); there has been continual interchange of ideas across ever-changing national borders (Hinrichs and Barnes, 2013); ideas have been developed and imported back to China (Elman, 2011).

It should be noted that these ideas are now part of the praxis of acupuncturists across the globe, far removed from Asia. As a result, concepts and theory are being reinterpreted and reconstructed in a global dialogue, rather than a specifically Asian context. The word 'traditional' is also potentially misleading. As mentioned below, the third category TCM is relatively new and developed under the influence of biomedicine and political pressure to be scientific (Birch and Felt, 1999, Scheid, 2002, Unschuld, 2010). As the subject is so complex and vast any given term will have its limitations.

The third category, TCM, refers to the most commonly known theoretical foundation of acupuncture, taught in modern Chinese universities and many western colleges. TCM developed in the 20th century through a process of standardisation that was applied to traditional praxis (Scheid, 2002, Unschuld, 2010). This process was influenced by biomedicine as well as wider social and political changes that were affecting Chinese society (Birch and Felt, 1999, Scheid, 2002, Unschuld, 2010).

The challenges of incorporating TEAM into scientific research are analysed in this chapter. Certain recent developments within TEAM occurred primarily in China, in these instances the term traditional Chinese medicine is used. The pilot study protocol was based on TCM ([chapter-VIII](#)). TEAM includes traditionally-based systems of acupuncture (TBSA) as well as other therapies such as herbal medicine and Qi Gong exercises. This project is primarily

⁴ The term 'herbal medicine' is in common usage to describe Chinese medicine decoctions, although they may consist of ingredients that are not herbs.

concerned with acupuncture. For simplicity, normally TBSA is used in this thesis rather than TEAM, even though many of the issues apply to TEAM in general and not just TBSA in particular.

I.3.2. Historical diversity and authenticity

In Chinese medical literature, ideas are often presented as a rediscovery of the true meaning of ancient texts. In Chinese thought, the citing of ancient texts carries with it a claim of authority and authenticity (Lloyd and Sivin, 2002). There were numerous factors within Chinese culture that led to *the authority of the past* such as: intellectual life being structured around coterie that saw themselves as protectors of a lineage; the dependency of the literati on the patronage of rulers; or the desire to avoid direct confrontation with contemporaries (Lloyd and Sivin, 2002). Unlike modern science where new knowledge replaces the old, 'in ancient China all knowledge of antiquity remained co-present, the older the more inherently valuable' (Lloyd and Sivin, 2002, p77). Thus, the citing of the historical texts often has a very different purpose within TEAM compared to biomedical textbooks. In biomedicine, an individual such as William Harvey may be cited as a point of interest but not with the intention that he should still be considered the definitive authority on blood circulation.

Claiming the authority of the past, engenders the belief that there is an authentic unitary theory of TEAM. Yet on closer investigation, an authentic unitary theory is a will-o'-the-wisp. Different generations, schools of thought, and individuals have been continuously assimilating and reconstructing TEAM, leaving a diverse and contradictory body of information (Hinrichs and Barnes, 2013, Scheid, 2002, Unschuld, 2010). Critically, even at the dawn of recorded Chinese medicine there was diversity. The key early text the Yellow Thearch's Internal Classic was almost certainly not written by one author and contains divergent opinions (Lloyd and Sivin, 2002, Unschuld, 2003).

Not only is there diversity but there is a considerable amount of synthesis and re-interpretation. Many books on acupuncture and moxibustion draw their inspiration from texts written in earlier ages (Shi and Zhang, 1998). For example, *The Great Compendium of Acupuncture and Moxibustion*, written in 1601, was in part a reinterpretation of *The Yellow Thearch's Internal Classic* (Shi and Zhang, 1998).

In the 20th century, traditional Chinese medicine was increasingly presented as being systematic - with a scientific theory (Hinrichs and Barnes, 2013, Scheid, 2002, Unschuld, 2010). It has been suggested that, quotes from the classics within traditional Chinese medicine textbooks were often shortened as a way of concealing the diversity and inconsistencies (Hinrichs and Barnes, 2013). It is through the modern textbooks, translated into English, that many practitioners in the UK have come to know and understand TEAM.

In summary we can note that, the praxis of TEAM has been diverse throughout history. Despite an enduring trait to refer to the authentic practice of the ancients, it is not possible to identify one style of praxis and realistically claim it has the privileged status of being 'authentic'.

I.3.3. Inter-style homogeneity and intra-style heterogeneity

Within TEAM acupuncturists often identify themselves as practising a particular style: Five-elements, Traditional Chinese Medicine (TCM), Stems and Branches are all considered essentially Chinese styles; there are other traditional styles such as Korean, Japanese and Toyohari. These styles with a documented historically based explanatory model can be identified as Traditionally-Based Systems of Acupuncture (TBSA) (Birch, 1997). Other classifications include Contemporary Styles of Acupuncture (CSA) and micro-systems (e.g. ear acupuncture) (Schnyer and Allen, 2002). In the UK the most common CSA is [western] medical acupuncture, whose practitioners state their practice is based on biomedicine (Hopton et al., 2012, White and Editorial-Board, 2009).

The styles of acupuncture are, however, only broad categories and there appears to be a synthetic relationship between them. Some colleges within the UK state that training is given in two different styles (Acupuncture Academy, 2018, CICM, 2018). These practitioners, for any given patient, may choose a Five-elements or a TCM approach (Winter, 2016). Practitioners, once trained, can go on to develop skills in other styles of practice, such as Japanese acupuncture (Fixler, 2018). Practitioners may incorporate ideas from a variety of other sources without undertaking formal training, such as other healing traditions or biomedicine. Synthesis is a distinct characteristic of TEAM, this was the case historically in China, it is part of praxis in modern China (Scheid, 2008), and currently in the UK.

The synthesis of ideas raises the question as to whether styles of practice can be delineated. As an example of how entangled different styles can be, consider the following: It might be supposed that to delineate medical acupuncture from TBSA would be relatively easy, considering it has been stated that medical acupuncture does not use traditional theory (White and Editorial-Board, 2009). Nevertheless, medical acupuncturists use the traditional acupuncture points (White et al., 2008). In recent decades, 'traditional' practitioners have started to use scalp acupuncture, a system of point selection based on the anatomical structure of the brain (Wang et al., 2007). In both situations, each style utilises knowledge derived from a seemingly antithetical style.

The issue of how reliably two styles can be differentiated is accompanied by a need to examine the degree of homogeneity within styles. A literature review, audit of case notes, and a practitioner survey found variability within the TCM acupuncture style for low back pain (Birch and Sherman, 1999, Sherman et al., 2001a, Sherman et al., 2001b). It has also been shown that TCM practitioners who saw the same patient made different diagnoses (Hogeboom et al., 2001). Thus, there is uncertainty whether two independent clinical trials that report using the same 'style' can be reliably considered as equivalent.

I.3.3.a. Defining styles used within research

Numerous controlled trials have used acupuncture protocols that did not reflect clinical practice and have lacked external validity (Stux and Birch, 2001). This led to recommendations that the style of acupuncture should be reported (Lao et al., 2001, MacPherson et al., 2010), and protocol development should follow rigorous processes: literature reviews, consultation with experts, practitioner and Delphi surveys (Alraek, 2014, Birch and Alraek, 2014, Hammerschlag and Stux, 2001, Schnyer and Allen, 2002). Given the synthetic relationship between styles, defining and delineating theoretical systems is a key challenge. Styles may use very different language and theoretical explanations, yet, use the same acupuncture points. Mapping the physical components of treatments is necessary to illuminate the homogeneity and heterogeneity of practice. Thus, the mapping of acupuncture in practice is a prerequisite for the evaluation of the external validity of clinical trials.

I.3.3.b. Styles of practice in a state of flux

TEAM has evolved under a whole range of social and environmental factors (Hinrichs and Barnes, 2013). Indeed, TCM style only developed in the 20th century in response to political and social forces. The perceived threat from biomedicine created a desire for a systemised and unified theoretical system (Hinrichs and Barnes, 2013, Unschuld, 2010). The treatment and theoretical explanation of the menopause has changed, influenced by a biomedical understanding of physiology (Scheid, 2007). The illnesses confronting practitioners have also changed, leading to developments in theory and practice (Hinrichs and Barnes, 2013). Modern illnesses, caused by lifestyle factors such as diet and stress, are very different from the common diseases of a century ago (Sapolsky, 2004). We should not be surprised if acupuncture has adapted to address the conditions of the modern world. At the level of the individual, acupuncturists' theoretical understanding and practice may evolve in light of numerous factors such as experience, in response to scientific evidence, or as a

result of further study. As the praxis of acupuncture is in a state of flux, this necessitates research to investigate the changes and the factors that drive them.

I.3.4. Moxibustion

In TBSA, needling is intimately related to the use of moxibustion. Moxibustion is the burning of an herb called moxa (*Mugwort*, *Artemisia vulgaris*) to warm specific parts of the body, including acupuncture points. The Chinese word translated as ‘acupuncture’ is *zhenjiu* 针灸. This actually refers to both the use of needles *zhen*-针 and moxibustion *jiu*-灸. There are a variety of ways in which moxibustion can be used, including warm needling, moxa box, moxa sticks and indirect moxa using ginger (Cheng, 1999). It could be argued that needling and moxibustion are two components of one therapy. There are also other methods, often considered adjunctive, such as massage, guasha or cupping. In the context of many clinical trials, needling has been presented as a standalone treatment modality. It was the absence moxibustion in the evidence base that initially prompted this project.

I.3.5. The praxis of TEAM

A better understanding and greater utilisation of TEAM has been advocated to support acupuncture research (Cardini et al., 2006, Langevin et al., 2011, Luo et al., 2015, Witt, 2011). Traditional methods of classification of diseases *bianbing*-辨病 and of diagnosis *bianzheng*-辨证 are explained in this section. The description presented below is the TCM approach, and may not apply to all TBSAs. TCM is the dominant style in Chinese hospitals and teaching institutions today, and is taught in many Western colleges. Furthermore, in keeping with the historical diversity described above, it should be noted that *bianzheng*-辨证 ‘pattern differentiation’ did not always hold the central place in practice it does today (Scheid, 2002).

I.3.5.a. Nosology - *Bianbing*

Chinese medicine compendiums of case studies are often collated according to the *bing*-病, which can be translated as illness or disease. The *bing* is the main

presenting sign or symptom. Headache, lower back pain, nausea, diarrhoea and insomnia are all typical examples. However, there is often no direct correlation between bing-illness and biomedical disease categories. The process of assessing which particular sign or symptom should be classified as the bing-illness is known as *bianbing*. The meaning of *bian* 辨 is to analyse/ recognise/ tell the difference (Gu, 2003). However, treatment is not guided by *bianbing* but by a second process *bianzheng* 辨证, often translated as pattern or syndrome differentiation. Xu Lingtai 徐灵胎 a famous doctor from the Qing dynasty stated: a bing-illness will have several differential patterns, furthermore, a differential pattern will be present in different diseases (Wu, 1999)

证之总者为之病，而一病总有数证

zheng zhi zongzhe wei zhi bing, er yibing zongyou shuzheng

And so, participants entering a trial with a certain bing-illness, may present with a number of differential patterns and should receive different treatments, accordingly. For this reason, the failure to use pattern differentiation within clinical trials has been criticised (Birch, 1997, Luo et al., 2015).

1.3.5.b. Diagnosis - bianzheng: a subjective process

Differential patterns can be considered subcategories of a bing-illness. It should be possible to accommodate these subcategories within clinical research, yet, there is a further complication. Chinese medicine posits, as each patient has a unique set of signs and symptoms then the differential pattern is unique. The differential patterns described within textbooks should be modified by the practitioner according to the actual presenting signs and symptoms of each individual.

Not only are individual patients' differential patterns unique but these patterns, for any given patient, can change over time. As the signs and symptoms fluctuate there will be an associated modification to the differential pattern. This implies potential changes to the acupuncture points used, and how they are stimulated, throughout the course of treatment. The requirement for an

individual pattern differentiation and the temporal adjustment both conflict with the requirements of controlled trials.

Pattern differentiation is fundamentally a subjective process. *Zheng*-証 ‘pattern’ - originally meant ‘expostulate with a superior’ or ‘test and verify’ (Gu, 2003). Its radical indicates that it is connected to language. Rather than being a concept that is directly observable, it is theoretical. It is the process of analysing *zhenghou* 症候⁵ and constructing a relationship between them using the theory of Chinese medicine. *Zhenghou* are observable changes due to illness that can be perceived by either the patient or the practitioner, this term encompasses both signs and symptoms. Many of these signs and symptoms are subjective (Alraek, 2014). Therefore, the subjectivity of pattern differentiation stems not only from its analytical process (the debate), but also from the assessment of the clinical information upon which it is based; namely, the signs and symptoms. This subjectivity means the same patient may be ‘diagnosed’ differently depending on the practitioner. Variations in the diagnosis of patients with low back pain have been shown in previous research (Hogeboom et al., 2001).

I.3.5.c. Effectiveness may depend on individual skill

In Chinese philosophy, ideas are not abstracted but are seen as being embodied within the teacher (Lloyd and Sivin, 2002). The notion that doctors should practice self-cultivation is widespread. On a practical level for acupuncturists, this includes the ability to manipulate Qi with the needle; *shoufa*-手法 (literally hand method). It is believed, that this capacity is not simply a mechanical process but also requires focusing of the mind, the yi [intention] or zhishen [will], whilst inserting or manipulating the needle (Liu, 2009). The ability to deliver acupuncture and stimulate Qi is a skill, similar to sport or playing a musical instrument. Consequently, the practical ability of the practitioner may also play a role in clinical outcomes.

⁵ The pronunciation of zheng-症 ‘symptom/sign’ is exactly the same as zheng-証 ‘pattern/syndrome’, however, the characters and meanings are different.

I.3.6. Summary

The traditional diagnostic process (pattern differentiation) presents three particular challenges for clinical research: individualisation, adaptation through the course of treatment, and subjectivity. Acupuncture protocols cannot meet the requirement of a controlled trial to be reproducible and at the same time have complete external validity. Even if one is able to clearly delineate a style and develop a protocol using rigorous methods ([section-I.3.3.a](#)) some compromises with external validity may still be necessary. Moreover, as the quality of the acupuncture delivered might depend on the skill of individual practitioners results may be difficult to generalise. [Box I-3](#) summarises the key challenges of TEAM that need to be addressed within the research framework ([chapter-II](#)). Methods of ameliorating these challenges are proposed in [sections-II.4/5](#).

Box-I.3: Challenges from the characteristics of TEAM

- The diversity in praxis
- The lack of a single 'authentic' theory
- Understand the heterogeneity/homogeneity within any given style
- Understand the heterogeneity/homogeneity between styles
- The subjective assessment of signs and symptoms
- The subjective process of pattern differentiation
- The skill-based elements of practice

I.4. Placebo

The first paradox of acupuncture research centres around the concept of placebo (box-I.1). Some commentators have stated that the evidence indicates acupuncture is a ‘theatrical placebo’ and the use of a ‘placebo’ raises ethical issues with regards to deceiving patients (Colquhoun and Novella, 2013, O’Connell et al., 2009). The relevance of efficacy and effectiveness data is not only a methodological discussion but also an ethical and political one (Witt, 2009). There is a perception that efficacy, comparison against placebo, demarcates legitimate from illegitimate medicine (Kaptchuk et al., 2009). Consequently, placebo controls are: required by key stakeholders (Bovey, 2016, Cummings, 2009b, Molsberger et al., 2006); described as essential in researching physical therapies (Fregni et al., 2010); established as the central dogma held by acupuncture detractors (Colquhoun and Novella, 2013, O’Connell et al., 2009, Vickers et al., 2013).

It has been suggested that by adopting a translational research approach, ‘a clear understanding will emerge of how to create a protocol that appropriately controls for non-specific effects of acupuncture by mimicking but not contributing to the specific acupuncture related therapeutic process’ (Langevin et al., 2011, p6). However, it may not be possible to ever develop a sham procedure (Birch, 2006). If this is the case, acupuncture appears to face an intractable problem, the requirement to be compared to a placebo when such a procedure is impossible. The crux of the matter, however, is not the placebo acupuncture procedures per se, it is the concept of placebo.

I.4.1. Definitions of placebo and placebo effect

The placebo effect is synonymous with placebo response; and is linguistically linked with placebo interventions. Influential definitions intrinsically connect the placebo response with a placebo intervention ‘the psychological or psycho-physiological effect produced by placebos’ (Shapiro, 1959). Whereas others do not require a placebo intervention, ‘a change in a patient’s illness attributable to

the symbolic import of a treatment rather than a specific pharmacologic or physiologic property' (De Craen et al., 1999, p513). It has been stressed that the definition of placebo effect should imply a causal, not temporal, relationship to a placebo intervention; it is not simply the change after the administration of a placebo. A temporal relationship includes regression to the mean and the natural course of disease, neither of which should be considered part of the placebo effect (Endres, 2008, Hróbjartsson, 2002). Placebo interventions have been defined as being inert/non-therapeutic (Abhishek and Doherty, 2013, De Craen et al., 1999, Hróbjartsson, 2002). This leads to an inherent contradiction with the presumption of a causal relationship. If an intervention is inert, it cannot *cause* any change; a placebo cannot *cause* a placebo effect (Moerman and Jonas, 2002). Therefore, despite the linguistic association these are two separate concepts.

Table-I.1: Definitions of the placebo effect/ response

The psychological or psychophysiological effect produced by placebos	Shapiro*
A change in a patient's illness attributable to the symbolic import of a treatment rather than a specific pharmacologic or physiologic property	Brody*
The difference in outcome between a placebo treated group and an untreated control group in an unbiased experiment	Gøtzsche*
An effect in which individually or culturally based expectations for a treatment cause are contributory to physical or psychological improvement after such a treatment	Anderson*
The symptomatic improvement experienced by a patient on receiving an intervention, or a set of interventions that are regarded as inert and non-therapeutic for the condition ('placebo') compared to those who receive no such intervention(s)	Abhishek and** Doherty

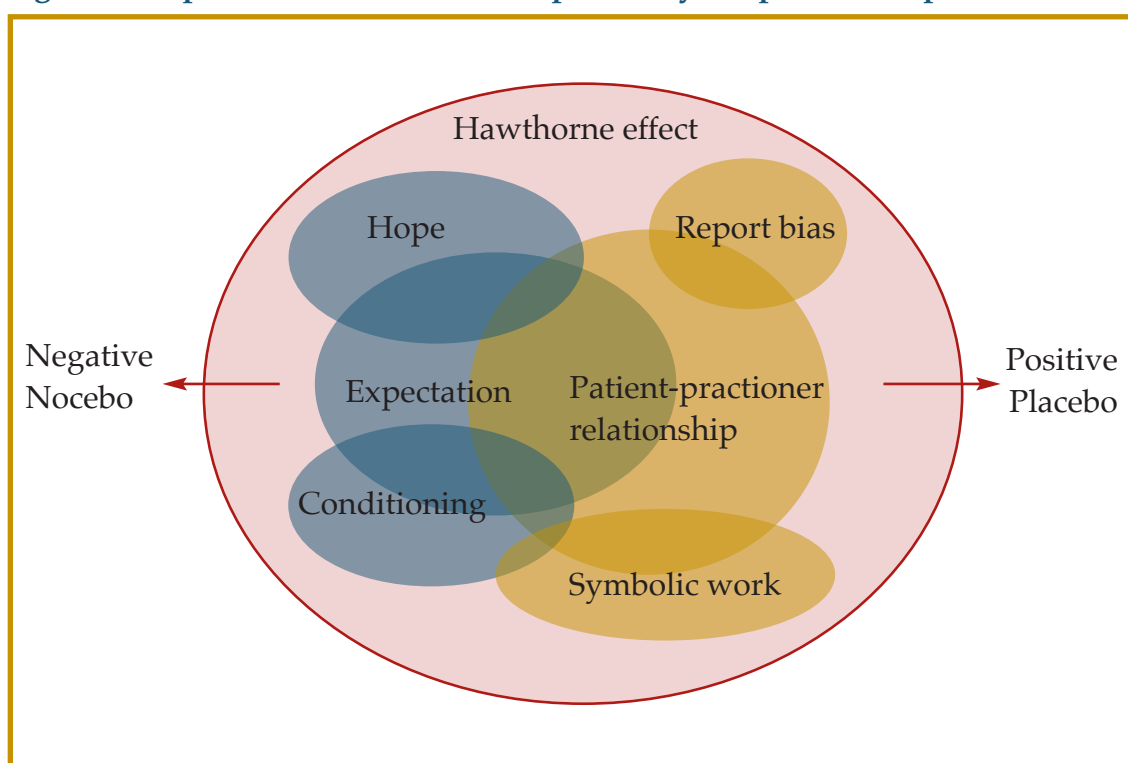
*Cited in De Craen et al (1999); **(Abhishek and Doherty, 2013)

1.4.1.a. Definitions of placebo effect/ response

Definitions of the placebo response indicate that it is perceived as a psychologically mediated effect (table-I.1). It is not a unitary phenomenon,

however, and a single theory of the placebo effect is an inadequate explanation. The changes ought to be conceptualised as being caused by a spectrum of factors - for example, expectation, hope, conditioning, anxiety reduction, report bias, symbolic work, narrative and embodiment (Grünbaum, 1986, Hróbjartsson, 2002, Kaptchuk et al., 2009). These factors do not necessarily have a positive effect. If the effect is negative it is known as nocebo. Figure-I.1 is a visual representation of some of the potential factors involved in the placebo response.

Figure-I.1: Spectrum of factors encompassed by the placebo response



1.4.1.b. Definitions of placebo intervention

There is no clear, logical and agreed definition of placebo (De Craen et al., 1999, Frenkel, 2008, Gotzsche, 1995, Hróbjartsson, 2002, Moerman and Jonas, 2002). Definitions often centre upon the notion of specificity; specific effects are associated with the 'real' treatment, non-specific related to placebo. However, there is no satisfactory explanation of the meaning of specific and non-specific (Grünbaum, 1986, Hróbjartsson, 2002). Non-specific has been used to refer to effects that are not unique to a given treatment. As a result, opiates and non-

steroid analgesics would be considered non-specific, as their effects are not unique, and should therefore be redefined as placebo. Another definition is that non-specific pertains to effects for which there exists no empirically supported theory. This would require redefining interventions that are known to be effective, such as lithium, as placebo (Hróbjartsson, 2002). As there is currently no clear understanding of the mechanism of acupuncture (Langevin et al., 2011, Napadow et al., 2008), acupuncture itself also should be categorised as placebo - not just the so-called sham procedures. Specific and non-specific can also mean non-psychological and psychological factors, respectively (Grünbaum, 1986, Haake et al., 2007).

Table-I.2: Definitions of placebo interventions

Definition of placebo interventions	
Any therapeutic procedure which has an effect on a patient, symptom, syndrome or disease, but which is objectively without specific activity for the condition being treated	Shapiro*
An intervention designed to simulate medical therapy, that at the time of use is believed not to be a specific therapy for the condition for which it is offered	Brody*
An empty preparation or intervention imitating an effective preparation or intervention where one must decide on the "emptiness" of the preparation in each particular situation	Hornung*
An intervention, or a set of interventions that are regarded as inert and non-therapeutic for the condition	Abhishek and Doherty**

*Cited in De Craen et al (1999); **(Abhishek and Doherty, 2013)

Some definitions of placebo avoid the terms specific and non-specific, for example Grünbaum uses *characteristic* and *incidental* factors (Grünbaum, 1986). Characteristic factors are those that according to a theory are considered responsible for the therapeutic effect. Incidental factors are not considered to be active (Grünbaum, 1986, Hróbjartsson, 2002).

Whilst there are differences in the terminology used, non-specific or incidental, the central problem is the same; the categorisation of an intervention as a placebo depends on an assumed theory. Thus, the classification process is

subjective. This is apparent in three out of four definitions of placebo set out in [table-I.2](#), as they contain expressions indicating subjective decisions; believed, one must decide, regarded. Shapiro attempts to confer the status of objectivity on the definition of placebo. It remains a subjective definition, however, because it requires the notion of an ‘inadvertent placebo’ - where a whole professional community believes an intervention to be placebo but are mistaken (Grünbaum, 1986). Subjectivity means identification of placebo interventions is particularly vulnerable to bias; conflicts of interest, preconceptions, and power structures ([section-II.2](#)). Therefore, the actual intervention should be described rather than using the potentially misleading term placebo.

I.4.2. Interpretive bias and placebo

The lack of agreed definitions for both placebo intervention and placebo response, and the false association between the two concepts, appears to be central to the paradoxes of acupuncture research ([box-I.1](#)). Different interpretations of the evidence-base hinge on the concept of placebo. The German acupuncture trials ([section-I.2](#)) often had three arms, real acupuncture, sham and conventional care. The results typically followed the same pattern; true acupuncture was superior to usual care, but compared to sham acupuncture the effects were relatively small (Cummings, 2009a, Vickers et al., 2012, Vickers et al., 2017). O’Connell et al (2009) disputed the interpretation of one of these trials, claiming it ‘offers an example of how, with interpretive bias, even gold standard research methodologies can lead to questionable conclusions’ (p405). The trial compared, real acupuncture, sham acupuncture, and multimodal conventional therapy (physiotherapy, drug, massage) for back pain (Haake et al., 2007). The following sections examine the arguments put forward by O’Connell et al (2009) as well as other authors.

I.4.2.a. Risk of bias stems from the subjective definitions

Some trials have used placebo procedures that appear to have a high risk of being too similar to acupuncture (Appleyard et al., 2014). There is a clear risk of

bias if procedures regarded as placebo are not in fact inert - inadvertent placebos ([section-I.4.1.b](#)). Yet, reviews have accepted any intervention described by the trial authors as placebo, apparently without any method of assessment (Madsen et al., 2009, Manheimer et al., 2010, Manheimer et al., 2007, Manyanga et al., 2014, Vickers et al., 2017, White and Cummings, 2015).

Remarkably similar procedures have been *regarded as* both placebo and acupuncture. The Streitberger needle was defined as a placebo because it does not puncture the skin, although it does apply pressure (Streitberger and Kleinhenz, 1998). A seasickness band is accepted as an active treatment as it applies pressure, although it does not puncture the skin (Lee et al., 2015). Superficial needling has been considered the active treatment one year and placebo the next, by the same reviewers (Ernst and White, 1998, White and Ernst, 1999). Ambiguities and contradictions such as these will inevitably lead to a lack of clarity within systematic reviews and the wider debate. A method of assessing the potential risk from the placebo procedures needs to be developed (Appleyard et al., 2014).

The lack of an objective definition of placebo acupuncture not only has a detrimental effect on acupuncture research but also potentially misleads scientific understanding of particular conditions. A Cochrane review concluded that placebo can influence patient reported outcomes, especially for pain. This assessment was primarily based on the German acupuncture trials, that combined low risk of bias with large effects (Hróbjartsson and Gøtzsche, 2010). Hróbjartsson and Gøtzsche were aware that active components might be retained in certain placebo interventions. They considered some trials to have an effect that was 'not related to the treatment ritual alone' and excluded them, such as movement techniques for post operative pain (Hróbjartsson and Gøtzsche, 2010, p5). They made a subjective assessment that sham acupuncture was 'only a ritual', and so inferred that pain responds to placebo. Other reviews also concluded that pain was amenable to placebo, but these assessments were

based on placebo-acupuncture, too (Abhishek and Doherty, 2013, Zhang et al., 2008). As placebo acupuncture retains potentially active components (Appleyard et al., 2014), the conclusion that pain is amenable to the spectrum of factors that constitute the placebo effect could be mistaken.

An estimate of a generic placebo effect, based on low quality evidence from unrelated interventions, has been used to downplay the potential benefits of acupuncture. Kavoussi (2015) argues that acupuncture is a placebo - based on acupuncture versus sham evidence. When considering whether the benefits outweigh the risks he cited a Cochrane review to suggest that ‘placebo treatments had only a modest effect on subjective outcomes such as pain relief’ (Kavoussi, 2015, p7). In that Cochrane review, as has been stated above, sham acupuncture had large effects (Hróbjartsson and Gøtzsche, 2010). Kavoussi chose to emphasise the estimate of a generic placebo effect, that combined the large effects of placebo acupuncture with the low or no effect of unrelated placebo interventions. There was no scientific reason to use the generic estimate, given that high-quality evidence relating specifically to sham acupuncture was available.

1.4.2.b. Risk of bias from the false notion of a causal link

In their analysis, O’Connell et al state that ‘there is a large literature of the variations in the magnitude of effect between different forms of placebo’ (O’Connell et al., 2009, p396). They imply that this justifies their interpretation that sham acupuncture out-performed conventional care due to the placebo effect. The argument relies on the inherently contradictory proposition that an inert placebo causes a placebo response. They cite a comparison between non-penetrating acupuncture and a sugar pill, (Kaptchuk et al., 2006), to establish the notion of ‘different forms of placebo’; along with a couple of other non-acupuncture examples. As the sham-acupuncture and sugar pill are both described as placebo the greater effect of sham-acupuncture is assumed to be a ‘strong’ placebo response. No consideration is given to the

possibility that the non-penetrating acupuncture may have an active physical component. Despite the fact that a similar procedure, seasickness band, has been classed as acupuncture (Cheong et al., 2013, Lee et al., 2015).

The 'different forms of placebo' hypothesis is then used to justify the explanation that minimal acupuncture out-performed conventional care, in Haake (2007), due to the placebo response. This justification is facilitated by linking non-penetrative with minimal acupuncture. These are not the same procedures but have both been labelled placebo. If we avoid the term placebo and describe the procedures, as well as indicate that they are inert (which they should be by definition [section-I.4.1.b](#)), then O'Connell et al's (2009) argument is as follows:

There are lots of examples of different forms of inert interventions that are more effective than other inert interventions (See Moerman (2002) for review). Of particular relevance, inert non-penetrating acupuncture is more effective than an inert sugar pill. Therefore, the inert minimal acupuncture is more effective than conventional care because its a strong inert intervention.

This clearly makes no sense, whatsoever. Readers are referred to Moerman (2002) for further examples of different forms of placebo. Moerman argues cogently against the causal link between placebo intervention and placebo response. Unfortunately, O'Connell et al (2009) do not appear to have understood this point. Instead they further entrench the causal link proposition by claiming placebo acupuncture *causes a strong* placebo effect. If placebos are defined as being inert, they cannot cause anything nor can there be different strengths of placebo.

I.4.2.c. Risk of bias from conflating the spectrum of effects

The placebo effect is not a unitary phenomenon, it can be mediated by a number of different factors: the Hawthorne effect - i.e. benefit derived from participating in a clinical trial; the patient-practitioner interaction; and the patient's response to a therapeutic ritual (Hróbjartsson, 2002, Kaptchuk et al.,

2008). O'Connell et al (2009) suggest that several factors could explain the Haake (2007) study: such as statistical regression, natural history, and the Hawthorne effect. Others also state that 'what is being seen is regression to the mean, which is very likely to be the main reason why acupuncture sometimes appears to work when it does not' (Colquhoun and Novella, 2013, p1362). However, regression to the mean and natural history should not be confused with the placebo effects (Endres, 2008, Hróbjartsson, 2002). In the Haake (2007) study, these two factors would have been equally present in all three groups, thus cannot explain the differences *between* groups. The conventional therapy group were given an active treatment with more contact time compared to the acupuncture groups (Endres et al., 2008). Therefore, the Hawthorne effect and patient practitioner interaction cannot explain the difference between groups, either. Essentially, these are all misleading arguments that were accounted for by the study design.

1.4.2.d. Possible mechanisms ritual and expectation

By abandoning the unitary concept of the placebo effect a more rational understanding of the phenomenon can emerge. As just described, some of the factors within the spectrum cannot explain the difference between acupuncture and conventional care. However, expectation and the ritual of acupuncture are legitimate possibilities.

Expectation

O'Connell et al (2009) assert that a novel treatment, such as acupuncture, would have an advantage when compared to a conventional treatment, which had already failed the participants and was likely to induce a nocebo effect (O'Connell et al., 2009). Expectation may play a role (Bishop and Lewith, 2008, Kalauokalani et al., 2001a, Linde et al., 2007), but there are reasons to suggest the situation is nuanced and complex. The rationale for positive expectations appears to rely on acupuncture being 'theatrical', but fails to consider that the associated philosophy may equally engender negative expectations. Indeed, participants in trials, when asked, did not express simple positive expectations

of acupuncture (Hsu et al., 2014, Kaptchuk et al., 2009). Brain imaging research suggests that expectancy induced placebo and acupuncture have different mechanisms (Kong et al., 2009a). Some of the evidence vis-à-vis expectation derives from conditioned expectancy, in healthy volunteers within short timeframes (Kong et al., 2009b). Caution should be exercised when using evidence based on conditioned expectancy to make inferences about chronic pain conditions and expectation. Evidence suggests that acupuncture maybe effective in animals (Habacher et al., 2006), it seems unlikely that this would be due to expectation or theatre.

The impact of expectation on clinical outcomes needs to be further investigated. As yet, according to recent systematic reviews, there is not enough evidence to draw firm conclusions (Colagiuri and Smith, 2012, Prady et al., 2015). The issue of the nocebo could be addressed by comparing acupuncture to conventional treatment for those who have not received any kind of intervention. Acupuncture may be even more effective when delivered early on in the progression of the disease.

Ritual - a specific effect

In discussing whether or not placebo should be used within mainstream medicine O'Connell et al (2009) feel that, this is not an unreasonable position provided that '...researchers are cautious and clear about whether a response is due to *non-specific* placebo effects, or *specific* to the treatment tested, and it is important to perform and interpret controlled trials with this distinction in mind...' [emphasis added] (p404). Because the concept of specificity is not clearly defined in relation to placebo ([section-I.4.1.b](#)), it is not possible to know exactly what O'Connell et al (2009) mean without further clarification. For example, it could mean that researchers ought to be clear that there is currently no theoretical explanation for the effects, or that the effects are due to psychological factors.

Greater clarity would be achieved by using the term 'specific' as typically defined in dictionaries: belonging or relating uniquely to particular

[component] subject (Oxford Dictionary of English, 2005). If acupuncture produces beneficial effects due to its 'theatrical ritual', that are not produced by other therapies, then these effects are specific - they cannot be non-specific effects. As the acupuncture ritual has some aspects, such as touch, that are shared with other therapies, it would be better to use Grünbaum's (1986) term 'characteristic' when referring to the treatment ritual, rather than 'specific' which implies uniqueness.

1.4.2.e. The devaluation of psychological effects

Opponents of acupuncture appear to harbour ontological assumptions about what ought to be considered real or unreal. Colquhoun and Novella state: 'It is not possible to say whether the observed difference is a *real* [emphasis added] physiological action or whether it is a placebo effect of a rather dramatic intervention' (2013, p1361). This contrasts placebo effects with what is 'real'; the implication is placebo effects are 'not real'. Placebo effects do not solely exist subjectively, in fact they are measurable and observable (Frenkel, 2008, Moerman, 2013). Examples of observable placebo effects include, disappearances of peptic ulcers, and specific changes in brain activity observed in patients with Parkinson's disease using positron emission tomography (PET) (Frenkel, 2008, Moerman, 2013). As both 'real' treatment and placebo treatment lead to observable changes in biomarkers then it is misleading to describe one as 'real' and imply the other is 'unreal.' The perception of real or unreal effects, naturally, has a bearing on how the results are interpreted.

There may be some psychologically mediated effects that are characteristic to acupuncture. There may also be incidental psychologically mediated effects, that acupuncture shares with other therapies. The assumption that psychologically mediated effects are not worthwhile is questionable, especially from the patients' perspective (Ortiz et al., 2016, Walach et al., 2006). If these psychologically mediated effects are conceptualised as being due to a 'meaning response' with an acknowledgement that this ultimately entails a physiological

effect, then these benefits will be seen in a more positive light, than if they are characterised as unreal placebo responses (Moerman, 2013).

1.4.2.f. Ethical issues of placebo

The use of placebo is an ethical problem in both practice and research. Placebos appear to be widely used in practice but remain controversial (Asai and Kadooka, 2013, Bostick et al., 2008, Hróbjartsson, 2008). There are a number of issues including, the breakdown in trust due to deception, and potential harm to future patients from over-prescription of antibiotics (Hróbjartsson, 2008). The ethical considerations have been cited as reasons for not recommending acupuncture in national guidelines (Bovey, 2016). Yet, many of these concerns do not apply to acupuncture. In practice, patients are obviously aware if they have acupuncture and so there is no deception. The issue of overprescription of medication is not relevant to the use of acupuncture. Acupuncture could even reduce the use of antibiotics as placebo. There are minimal harms associated with acupuncture (MacPherson et al., 2001, Witt et al., 2009). Ethical considerations regarding the use of acupuncture should be based solely on issues that pertain to acupuncture.

The German acupuncture trials informed participants that they might receive one of two active treatments. This has been described as ‘falsely’ informing participants that they would receive an active treatment, when they might receive a placebo (Hróbjartsson and Gøtzsche, 2010). Hróbjartsson and Gøtzsche (2010) recommend that participants *should* be informed that they might receive a placebo. This would require researchers to deceive participants about the current state of scientific knowledge.

The thought of being allocated to sham acupuncture can cause anxiety for participants in trials; concerned that any response implies gullibility (Bishop et al., 2012a, Bishop et al., 2012b, Kaptchuk et al., 2009). Consequently those who respond to sham acupuncture may suffer as a result believing the effects were purely psychological (Bishop et al., 2012a). Researchers should consider

whether using the term 'placebo' in acupuncture research is ethical. It has been suggested, this could be ameliorated by providing patients with a more positive view of placebo. Although, it may be necessary to address perceptions of placebo in the wider society, also (Bishop et al., 2012b). There is no clear definition of placebo ([section-I.4.1](#)), so it seems likely that some confusion will remain among participants.

I.4.3. Separation of psychological effects from acupuncture

Traditional ideas understand acupuncture as a process of stimulating Qi to flow around the body. A passage in the Yellow Thearch's Internal Classic - Chapter 67: The Application of the Needle - suggests that stimulation of Qi is not limited to the time of needle insertion ([box-I.4](#)) The passage is a conversation between the Thearch and his minister Qi Bo, the standard structure of the text.

Box-I.4: The Application of the Needle - Yellow Thearch's Internal Classic

The Yellow Thearch asked Qi Bo:

I have been informed by you of the nine needles, and how they are applied to treat the people. However, all the people have different physical appearances of blood and Qi.

It happens that [a patients] spirit is excited, and the Qi move **even before the needle** have been applied.

It happens that the Qi and the needles confront each other.

It happens that the needle was withdrawn, but the Qi flow by themselves.

Translation by Unschuld (2016)

The second and third scenarios are: that Qi moves whilst the needle is inserted, (confront each other); that the sensation remains once the needle has been withdrawn. However, it is the first situation described that is most salient. When the spirit is excited, the Qi begins to move prior to the needle being inserted. Bringing a needle close to the skin stimulates the Qi, possibly as the patient anticipates the event. It is the movement of Qi that is considered to be therapeutic, not simply the insertion of a needle. The Qi can move before,

during and after needle insertion. In other words, TBSA theory conceptualises awareness/sensation as part of the mechanism of acupuncture. Therefore, procedures which lead participants to believe a needle is going to be inserted - which any placebo acupuncture must do - should be considered active in terms of TBSA theory.

The notion of 'psychological factors' has a broad meaning in the context of placebo. It encompasses more than something that one thinks or believes - but also the awareness of, and subconscious responses to, what is happening. It has been suggested that placebo is less about cognitive processes but more an experiential bodily engagement (Frenkel, 2008). Increased awareness of certain parts of the body may amplify local analgesic responses (Benedetti et al., 1999, Liu, 2009). Indeed, this kind of spatially directed awareness may be related to the mind-body exercises of TEAM; Qi Gong and Taiji Quan. For some practitioners, combining these forms of exercises with the acupuncture treatment is a part of everyday practice.

There are a number of biomarkers that correlate with needling; endogenous opioids, immune system markers, cardiovascular activity, gastrointestinal function and fMRI detected brain activity (MacPherson and Hammerschlag, 2012). However, biomarkers that correlate with sham acupuncture are also associated with 'placebos' such as creams or saline injections (Kaptchuk et al., 2010). This can 'blur the distinction between a "real" treatment with specific action mechanisms and the placebo treatment without such "specific" mechanisms' (Endres, 2008, p1012). It should, perhaps, be expected that the biomarkers associated with acupuncture and psychological effects overlap, because both must depend on endogenous systems. It has also been shown that pharmacological and psychological components are sometimes inseparable (Petrovic et al., 2002). Whilst some overlap should be expected it would be premature to conclude that acupuncture effects cannot be differentiated from psychologically mediated effects. Differences in the mechanisms have been

observed between real acupuncture and sham acupuncture (Dhond et al., 2008a, Dhond et al., 2008b, Harris et al., 2009, Huang et al., 2012, Kong et al., 2009b).

I.4.4. Are any acupuncture trials placebo-controlled?

Given the ambiguity surrounding the term placebo, it is reasonable to ask whether any acupuncture clinical trials can be considered ‘placebo-controlled’. There are two ways in which we could define a placebo-controlled trial. The first is that participants in one of the arms receive a placebo intervention. The second is that the spectrum of factors that constitute the placebo effect are balanced between at least two arms of the trial.

Under the first definition there is no objective answer. As there is no agreed definition of the placebo intervention, it is ultimately a subjective assessment as to whether any clinical trial can be described as placebo-controlled. If we accept a placebo intervention should be inert (Abhishek and Doherty, 2013), and none of the control procedures described as sham or placebo can be considered inert (Appleyard et al., 2014, Zhang et al., 2015, Zhang et al., 2012), then there are no placebo-controlled acupuncture trials.

Using the second definition, some clinical trials can be considered placebo-controlled. Let us first consider the non-insertion method of placebo acupuncture control. This method typically uses specially designed needles that do not penetrate skin (Streitberger and Kleinhenz, 1998). In well-designed studies, where group allocation is successfully concealed, the only difference between groups is the insertion or otherwise of the needle. In other words, various factors that comprise the ‘placebo-effect’ ([section-I.4.1.a](#)) are balanced between groups. This type of control procedure tests the *specific effect of one component of acupuncture*: needle insertion. Similarly, well-designed studies that use shallow insertion at non-classical acupuncture points (minimal acupuncture) can balance the placebo effect between groups. Because this form of placebo acupuncture manipulates two variables, point location and the depth of insertion, interpretation of the results is less straightforward. It is not possible

to judge whether any differences are due to the change in point location or the change in depth of insertion. These trials cannot show a *specific* effect. If the trials had been designed to test either point location or depth of insertion, then inferences about the *specific* effects of these variables could be made.

As trials that use minimal acupuncture do not test for a *specific* effect it is debatable whether they should be considered placebo-controlled to the same extent as the non-insertion method. However, both of these methods manipulate only physical variables and psychological factors are balanced between groups. There are other so-called placebo procedures, such as when an acupuncture point is compared to either a non-acupuncture point or a point not thought effective for a particular condition (Appleyard et al., 2014). In well-designed trials, psychological factors can also be balanced using these methods. *Therefore, psychologically mediated components, which includes the acupuncture ritual and more broadly the placebo response, are not responsible for any observed benefits of the real acupuncture.* There is, however, one caveat. In most cases the acupuncturist must be aware of group allocation. There is a risk that the practitioner shows additional care and attention to the experimental group. The extent to which this is likely to be the case will depend upon the trial design. The effects can be minimised by reducing the amount of interaction between acupuncturist and participants.

I.4.5. Should the term placebo continue to be used?

O'Connell et al ask 'If the authors believed a priori that the sham group may be a plausible therapy, then why was this condition chosen and why call it a "sham"?' (2009, p400). It is a legitimate question. The answer appears to lie within the political sphere of research. Perhaps, the various procedures are commonly referred to as 'placebo' or 'sham' to meet the demands of gatekeepers and the expectations of the wider scientific community (Bovey, 2016, Colquhoun and Novella, 2013, Cummings, 2009b, Kaptchuk et al., 2009, Molsberger et al., 2006, O'Connell et al., 2009).

In the context of drug trials, a placebo pill appears to be a clear, logical and stable concept. A useful tool to help delineate the effects of a putatively active pharmacological ingredient from psychologically mediated effects. In the context of acupuncture, however, the concept of placebo is confused, self-contradictory and illogical. As outlined above, the definitions of a placebo intervention rely on subjective judgement. This has allowed markedly similar procedures to be classed as both real acupuncture and placebo. The self-contradictory proposition that an inert procedure causes an effect has been expanded to include the notion of 'strong' placebo, in an attempt to find an explanation for paradox 1 (box-I.1). The idea that some inert procedures are stronger than others is illogical.

Whilst the concept of a 'placebo-intervention' for acupuncture is highly problematic, this does not mean that research which controls for the placebo-effect cannot be conducted. In the 1990s it had already been argued that the control procedures were actually testing certain parameters of acupuncture, rather than acupuncture versus inactive placebo (Lewith and Vincent, 1995, Vincent and Lewith, 1995). However, this argument is obscured by the confused semantics. Description of trials that control for the placebo effect as placebo-controlled commonly leads to the assumption that an inactive placebo intervention was used. Greater clarity would be obtained by using the term psychologically mediated factors.

The use of the term placebo is perhaps the most commonly applied within research involving human participants. Consideration should be given to whether or not its continued use is ethical. Categorically stating that a particular acupuncture control is a placebo cannot be objectivity justified. Yet, allocation to the placebo-controlled group potentially causes harm to participants.

In conclusion the term placebo should no longer be used in acupuncture research. Instead, the actual intervention should be described and the so-called placebo response replaced by *psychologically mediated factors*. This would clarify

one of the central questions of acupuncture research: are the effects mediated psychologically or via the physical process, or a combination of both? The psychologically and physically mediated factors ultimately may not be completely separate. As the concept of placebo is confused, self-contradictory and illogical it is fundamentally unscientific. Therefore, it is ironic that those who prioritise placebo often claim to be defending scientific principles.

I.5. Aims of the research

The original inspiration for this project is to test whether the additional use of moxibustion would lead to greater clinical benefits. This requires the development a protocol for a randomised controlled trial. The methodological challenges of clinical research have been discussed in this chapter. The following chapter proposes a number of solutions to these challenges and establishes a research framework. The subsequent phases of this project are situated within the framework. Therefore, the aims and objectives are presented in [chapter-II](#) after the research framework has first been explained.

II. Methodology: a novel framework for acupuncture research

方
法
论

II.1. Introduction

The original inspiration for this project was to test the hypothesis that the additional use of moxibustion is clinically beneficial for painful conditions. This requires conducting an experiment to produce empirical evidence. But the characteristics of TEAM would appear to suggest it is incompatible with experimental research ([section-I.3](#)). According to the dominant view of research methodology, placebo-controlled trials are essential. However, placebo does not have a clear definition and it is highly unlikely that a placebo acupuncture intervention will ever be developed ([section-I.4](#)). These are the key challenges of acupuncture research discussed in [chapter-I](#). Solutions to these challenges are proposed in this chapter.

Quantitative methodology has been characterised by a hierarchical framework of methods, with reviews of placebo-controlled trials at the apex. However, alternative non-hierarchical and translational methodologies that have been suggested in the literature for acupuncture ([section-II.3](#)). Building on these approaches, a new research framework is developed in [section-II.4](#). The new framework includes a method of accommodating TEAM clinical reasoning, pattern differentiation, within clinical research ([section-II.4.5](#)). It also identifies how trials can be designed to investigate the non-psychological components of acupuncture, and thereby provide evidence that acupuncture is not mediated through ritual alone. In other words demonstrate it is not ‘just placebo’. The framework underpins the rationale for the subsequent parts of this project; an overview is presented in [section-II.5](#).

Scientific investigation is a social process and there is a perceived conflict between ‘science’ and complementary and alternative medicine (CAM) (Goldacre, 2008, Singh and Ernst, 2008). Acupuncture is part of this wider epistemological discussion regarding paradigms, pseudoscience and scientific truth. These issues influence the way acupuncture research is conducted and how results are interpreted. The researcher’s views are set out in the following

section. Within this project TEAM is considered 'knowledge' and an important element of a translational research approach. The justification for this assumption is presented in [section-II.4.6.a](#).

II.2. Epistemology

Within the thesis, for the most part, I have written in the passive voice to support myself to be objective. Nevertheless, I think it is important to acknowledge my objectivity is limited by biases, too. Therefore, in this section, epistemology, I have chosen to use the first person.

II.2.1. The relationship between Science and TEAM

In the UK, acupuncture is considered a CAM. Within the public discourse, CAM are often portrayed as 'other' or in opposition to biomedicine, and diverse therapies are implicitly associated with each other. In this section, I will discuss two epistemological approaches to the relationship between science and CAM: Epistemic relativism and the demarcation of science and pseudo-science.

II.2.1.a. *Epistemic relativism*

Some practitioners feel TEAM is fundamentally incompatible with science (Quah, 2003). The term 'paradigm', coined by Kuhn, is used to describe a world view - including its theory, textbooks, apparatus and methodology. Kuhn's ideas are thought to have given rise to relativism, the theory that there are no truths that are independent of human beliefs (Rosenberg, 2005). Paradigms can be considered incommensurable, and consequently science cannot be used to test a CAM theory. This notion of separate paradigms has been adopted by some supporters of CAM, enabling them to claim that the theory of their practice is of equal validity to science. The epistemology of this investigation disagrees with the doctrine of epistemic relativism. On the contrary, it sees the synthesis of modern science and TEAM as mutually beneficial.

II.2.1.b. The demarcation of science and pseudo-science

CAM are often described as pseudo-scientific, anti-scientific or pre-scientific (Colquhoun, 2018, Goldacre, 2008, Singh and Ernst, 2008). These are essentially negative characterisations, suggesting there is an antagonism between CAM and science. Pseudo-science is by definition something that is unworthy of proper scientific investigation; it is a derogatory term (Hansson, 2014, Rosenberg, 2005). Being pseudo-scientific is an accusation that can be levelled at any individual or organisation that is connected with acupuncture research regardless of credentials; such as Harvard University (Salzberg, 2009), New England Journal of Medicine (Novella, 2010), or the Washington Post (Salzberg, 2016). Opponents of CAM not only use terms such as pseudo-science but also claim to be defending scientific principles (Colquhoun, 2018, Goldacre, 2008, Novella, 2010, Salzberg, 2016, Singh and Ernst, 2008)

There is, of course, a fundamental question that needs to be considered before acupuncture can be described as pseudo or pre-scientific. What is science? I do not believe there is a simple answer to this question. Moreover, for it to be reduced to a brief discussion within this thesis would be ultimately misleading. It would also be necessary to examine in detail the relationship between science and TEAM. The interested reader can find more information on these issues elsewhere (Pickering, 2010, Scheid, 2002, Scheid and MacPherson, 2012, Unschuld, 2009, Unschuld, 2010). However, it is necessary to examine how science is represented within the CAM controversy, and how science is demarcated from pseudo-science. I believe these factors shape the design and interpretation of acupuncture research.

There are a number of different ways in which science has been demarcated from pseudo-science. One method is 'falsification', proposed by Popper (Hansson, 2014, Popper, 2002). For a theory to be scientific it must be falsifiable. Within this context, 'a theory is a set of interrelated constructs (variables), definitions, and propositions that presents a systematic review of phenomena by specifying relations among variables, with the purpose of explaining natural

phenomena’ (Kerlinger, 1973, p9). Typically, research seeks to verify a theory by examining hypotheses derived from it (Creswell, 2008). If experimental findings disprove one part of a theory, then the whole theory is incorrect (falsified) or at least requires modification.

A restricted vision of science-as-knowledge where scientists are ‘disembodied intellects making knowledge in the field of facts and observations’ remains influential to this day (Pickering, 2010, p49). I believe this vision is prevalent in the discourse on acupuncture research. Combined with the belief that science is demarcated from pseudo-science by falsification, it leads to the view that knowledge can be generated by the following stepped method:

- ▶ experts of TEAM refine TBSA to produce a single authorised theory
- ▶ the theory is then tested by dispassionate and objective scientists
- ▶ at some point a final verdict will be given on a single question. Does acupuncture work?
- ▶ if acupuncture fails the test, the theory is falsified and labelled pseudo-science.

The epistemology of this stepped method contains erroneous assumptions. In the first instance, the process of science is far more complex (Pickering, 2010). The scientific ideal may strive for objectivity, however, the practice of science is inevitably a social endeavour. Which means that factors such as, conflicts of interest, preconceptions, misapprehensions, as well as power structures, play a role in defining what gets researched and how results are interpreted. There are multiple ways in which subjectivity can affect the interpretation of data (Kaptchuk, 2003). Indeed, it has been suggested the recent interpretations of acupuncture research are biased (Birch et al., 2016). [Section-I.4](#) examines the subjective and self-contradictory definitions of placebo intervention and placebo effect. I believe the dogmatic attachment to placebo-controlled trials for acupuncture owes much to the social process of science and has little to do with rational scientific argument. An epistemology that fails to acknowledge the potential for subjectivity will not be able to investigate TEAM effectively.

The stepped method (above) also makes a misapprehension regarding theory within TEAM. As TEAM is such a diverse body of knowledge it cannot be described as systematic, moreover, it is not internally consistent ([section-I.3](#)). Therefore, using Kerlinger's definition, TEAM cannot be considered a scientific theory. This leads to an axiom of the research framework for acupuncture. The diversity of acupuncture practice, the synthesis between styles, and the absence of any one particular style that can justifiably be considered 'authentic' ([section-I.3.2](#)) mean:

There is no single theory of TEAM or acupuncture that can be falsified by means of scientific investigation.

It is perfectly possible that some parts of TEAM/TBSA are true, whilst others are not. Given the extent and breadth of TEAM, it would seem that the diametrically opposed standpoints, that it is either completely without validity or entirely correct, are equally implausible. Components of acupuncture practice can be investigated to produce empirical evidence. For example, the use of non-penetrating needles investigates the specific component of needle insertion ([section-I.4.4](#)). When the results of these specific tests are used to make generalised claims that 'the entire philosophy of acupuncture falls apart' (Singh and Ernst, 2008), the evidence is being misinterpreted ([section-II.4.3](#)). This is the fundamental flaw in the proposition of demarcating TBSA as a pseudo-science. It extrapolates experiments that test specific components of acupuncture to make generalised claims about a non-existent unified theory.

[Section-I.3](#) sets out the challenge of researching TEAM due to the diversity and plurality of praxis. It is worth noting that science and biomedicine also share these characteristics to some extent (Pickering, 2010, Scheid, 2002). Newton is considered one of the towering figures of science, yet, his ideas regarding alchemy are false. This does not invalidate his contribution to mechanics (Rosenberg, 2005). Therefore, we should not infer that the entirety of the classical text is invalid simply because one section is shown to be incorrect.

Negative results for one particular pharmaceutical do not invalidate the results of another. Similarly, if evidence suggests acupuncture is ineffective for condition X, this should not be used infer it cannot be used the condition Y.

The epistemology of this project is positivism. However, a positivist approach can only be applied to acupuncture if there is an understanding of its components. Hence, the title of the thesis states it is a 'component analysis approach'.

II.2.1.c. The ongoing re-evaluation of TEAM

The stepped method of investigating acupuncture outlined above carries with it an explicit notion that ideas are drawn from TEAM and then tested using scientific methods. It might be assumed that a component analysis approach would follow the same chronology. However, I would like to emphasise that in my view the relationship is more complex. TEAM is not a static entity, it is constantly being re-evaluated and has already been influenced by biomedicine (Scheid, 2002, Scheid and MacPherson, 2012). It can, and should, be influenced by insights from scientific investigation. Ultimately, one of the objectives of applying scientific investigation to TEAM should be to improve practice. This inevitably means that certain components could be rejected whilst others may be used to a greater extent.

Complex social processes influence not only science in general but also medicine and TEAM, in particular (Scheid, 2002, Unschuld, 2009, Unschuld, 2010) The practice of TEAM and the classical texts are continually being re-evaluated within the profession. The most prominent style within TEAM is traditional Chinese medicine (TCM). The development of TCM was influenced by the social and political atmosphere in China, in the latter half of the 20th century, then moulded to fit the expectations of Westerners (Hinrichs and Barnes, 2013, Scheid, 2002, Unschuld, 2009). It is important to acknowledge these social factors and employ rigorous academic research to better understand TEAM and how it has developed through the centuries.

As an example consider Qi, perhaps the best-known concept of TEAM. Detractors claim there is no evidence for the existence of Qi, and cite this as a reason for considering acupuncture to be pseudo-science (Singh and Ernst, 2008). The implicit assumptions of this claim are: Qi is a single entity; and it must be separate from known phenomena, such as electromagnetism or biochemicals. This is a misconception of Qi. A misconception that may partly derive from Qi being translated as 'energy' or 'life force' by practitioners and advocates of acupuncture. These translations were adopted to meet expectations in the West and do not reflect the original Chinese meaning (Unschuld, 2009). Using the term 'energy' is especially unhelpful as scientifically there are clear definitions of energy. The philosopher Xunzi explicitly contrasts Qi with life:

Fire and water have qi but do not have life. Grasses and trees have life but do not have perceptivity. Fowl and beasts have perceptivity but do not have yi (sense of right and wrong, duty, justice). Men have qi, life, perceptivity, and yi.

水火有气而无生,草木有生而无知,禽兽有知而无义人有气,有生,有知,亦且有义。9:69/127 (Sturgeon, 2018)

Qi is polysemous, it has different meanings dependent on context. One translation is $E=mc^2$ (Zhang, 2002), in short Qi refers to everything. Using this definition it is nonsensical to claim that science cannot demonstrate its existence. The modern Chinese word for anger *shengqi* 生气 could be literally translated as growing Qi; again, anger clearly exists. We are all familiar with the sensation and understand expressions such as 'blowing his top'. The word Qi can be used in a meaningful way to describe the experience of anger. The original concept Qi in relation to acupuncture was something that was experienced and cultivated, not an external phenomena (Birch, 2015b).

The concepts and theories of TEAM may in time be superseded by a biomedical understanding of phenomenon. Anger can be described in terms of physiological changes, to some extent (Garfinkel et al., 2016). A complete description of anger in terms of biomedical physiology would not render Qi

meaningless. Newtonian mechanics and understanding of gravity have been superseded by Einstein's theory, so have been shown to be false (Rosenberg, 2005). Nevertheless, Newtonian mechanics still has its uses in engineering. Similarly the reduction of pain by using acupuncture may, one day, be explained in terms other than the 'unblocking of Qi' ([section-III.4.2](#)). This does not mean that the original notion of blocked Qi won't remain useful in clinical practice.

It is not my intention to provide a comprehensive explanation of Qi, rather emphasise there is no clear-cut easy definition. In truth, scientific research into Qi is still in its infancy (Birch, 2015a). Claims about the existence or nature of Qi based on questionable translations are inherently flawed. Scientific research into TEAM and its concepts, such as Qi, needs to be conducted in tandem with historical and linguistic scholarship.

II.2.2. Rationale for focussing on the physical components

The overall project is concerned with understanding acupuncture in terms of its physical components. This pertains to what I believe is the fundamental question at the heart the placebo controversy ([section-I.4](#)): whether the physical components of acupuncture effect clinical outcomes. In other words, do variables such as the selection of points or depth of needle insertion have specific effects? Qualitative research into the theoretical ideas of practitioners, the meaning(s) of Qi and other concepts, are necessary and important research objectives. Pragmatic clinical trials and the experiences of patients are also important. However, these lie outside the scope of this project. Demonstrating that the physical components have a physiological effect, independent of any psychological factors, necessitates a primarily quantitative approach.

Practitioners of different styles may articulate their clinical reasoning using different theories and language ([section-I.3](#)). Potentially, the competing theories could be a distraction when analysing acupuncture in terms of the physical variables. If the acupuncture points used and needling methods are similar,

then should two styles actually be considered different? Therefore, to provide clarity, each phase of this project focuses on the physical components.

I believe that evidence that demonstrates acupuncture has specific physically mediated effects is vital for the wider acceptance acupuncture. In addition, trials that investigate specific parameters of acupuncture can be more readily linked to mechanism research (Lewith and Vincent, 1995). These trials can also improve understanding of what constitutes best practice. This may enable some of the disputes between styles of practice to be resolved.

II.3. Acupuncture research frameworks

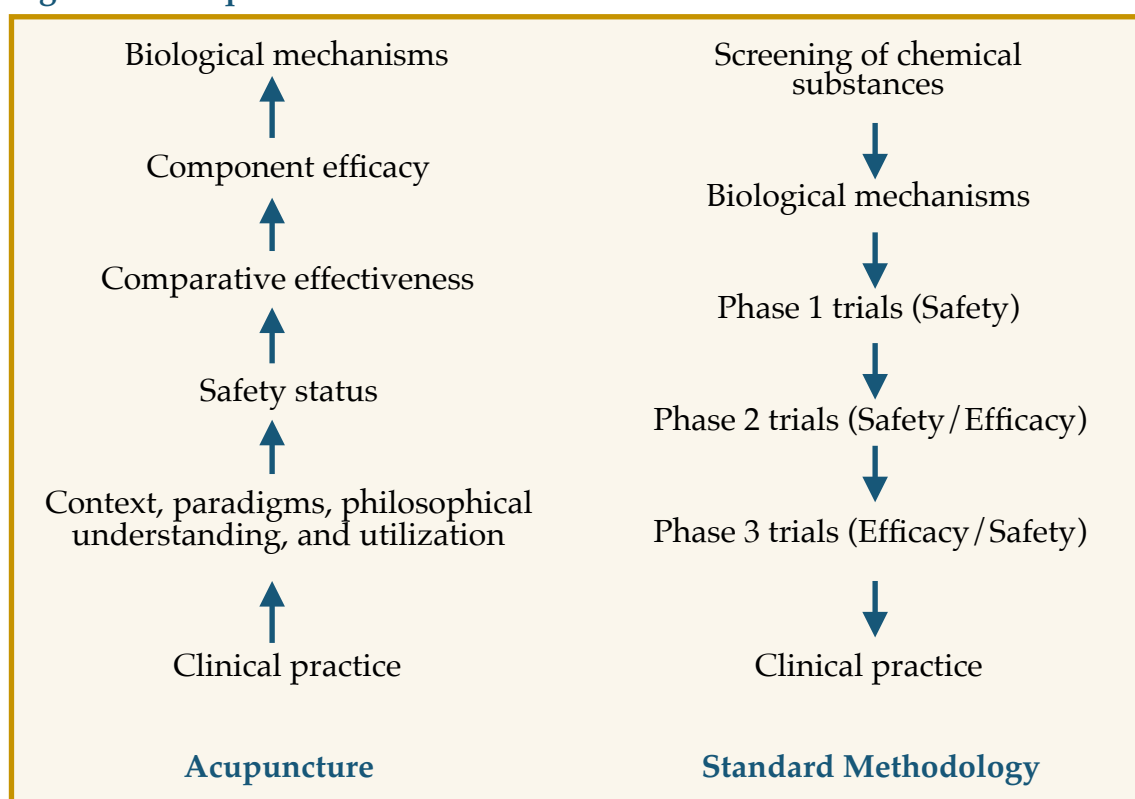
II.3.1. Alternatives to the hierarchical methodology

Acupuncture is a complex intervention as it involves not only the insertion of needles but also, moxibustion, other adjunctive methods such as cupping, and lifestyle advice (Langevin et al., 2011). Even if it is conceptualised as only the insertion of needles, it is still complex. Complex interventions can be defined by a number of dimensions, one of which is the degree of flexibility or tailoring permitted by the intervention (Craig et al., 2014). Adjusting the treatment according to an individuals presenting signs of symptoms is a key characteristic of TBSA (section-I.3.5). Heterogeneity is a particular challenge when evaluating complex interventions (Craig et al., 2014, Herbert and Bø, 2005). There is considerable diversity in the practice of acupuncture (section-I.3). For complex interventions, in general, it is recommended that research should take a more flexible, less linear, approach (Craig et al., 2014).

The standard methodological reasoning of evidence-based medicine posits a hierarchy of methods, with placebo controlled RCTs central to this approach (Cardini et al., 2006, Fønnebø et al., 2007, Walach et al., 2006). The first step requires a biomedical theory that suggests the intervention may work. Then RCTs should be conducted to demonstrate efficacy. Finally the intervention should be tested in a real-world setting. It has been suggested that for

acupuncture this process should be reversed from the norm, with a starting point in clinical practice and an endpoint of biological mechanisms ([figure-II.1](#)) (Cardini et al., 2006, Fønnebø et al., 2007). As acupuncture has been widely used in humans for centuries this indicates, but does not guarantee, the effectiveness and safety of the treatment (Cardini et al., 2006). Pragmatic trials should be conducted first, to establish whether acupuncture in its entirety has a beneficial effect and the procedures are safe. Importantly acupuncture has not developed from a biomedical theory, this is one of the reasons that has prevented the development of a placebo control ([section-I.4.](#)) Therefore, there can be no straight-forward efficacy trials and the methodological sequence has to be different.

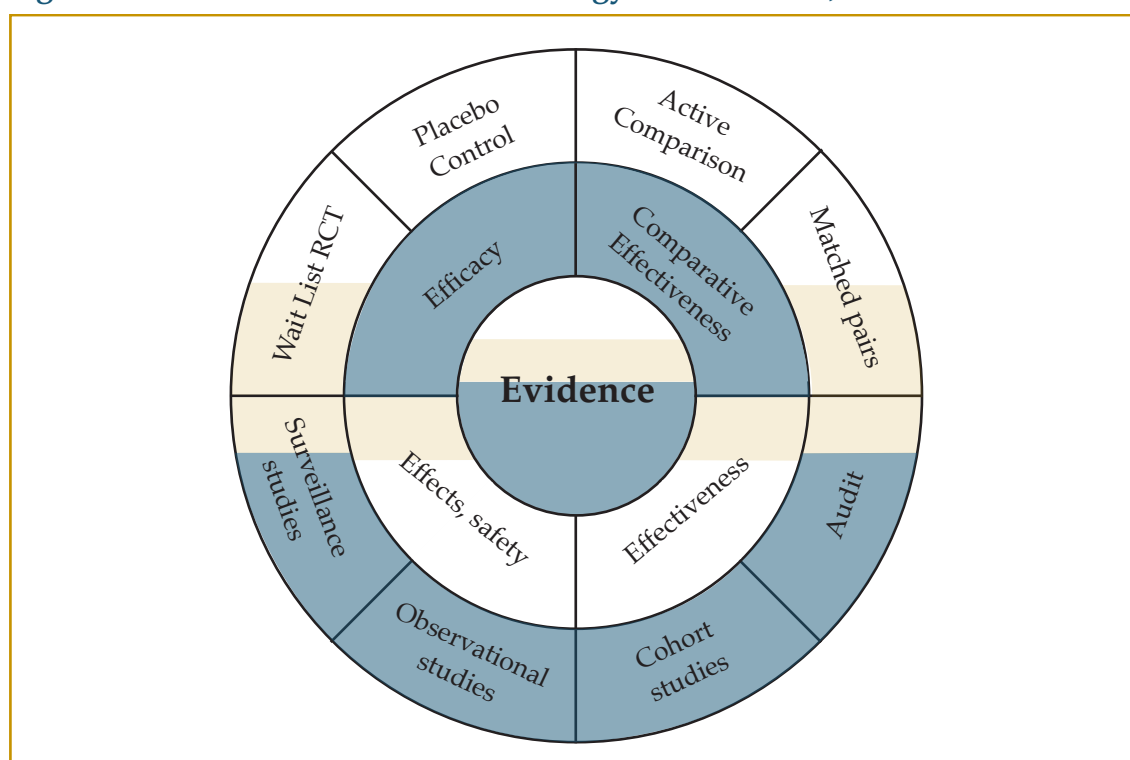
Figure-II.1: Sequence of research Fønnebø et al 2007



A circular approach has also been suggested, as every research method has strengths and weaknesses which cannot be resolved within that method itself (Walach et al., 2006). Triangulating results from different methods could provide a more powerful and comprehensive approach to evidence based medicine, compared to the prevailing hierarchical methodology (Walach et al., 2006). The

circular approach is illustrated in figure-II.2. Experimental methods that test specifically for efficacy are in the upper half of the circle. These trials will often lack external validity so need to be complemented by observational, non-experimental methods - the lower half of the circle - which are more descriptive in nature and describe real-life effects and applicability. Shading indicates the complementarity of experimental and quasi-experimental methods, of internal and external validity.

Figure-II.2: Circular research methodology (Walach et al., 2006)



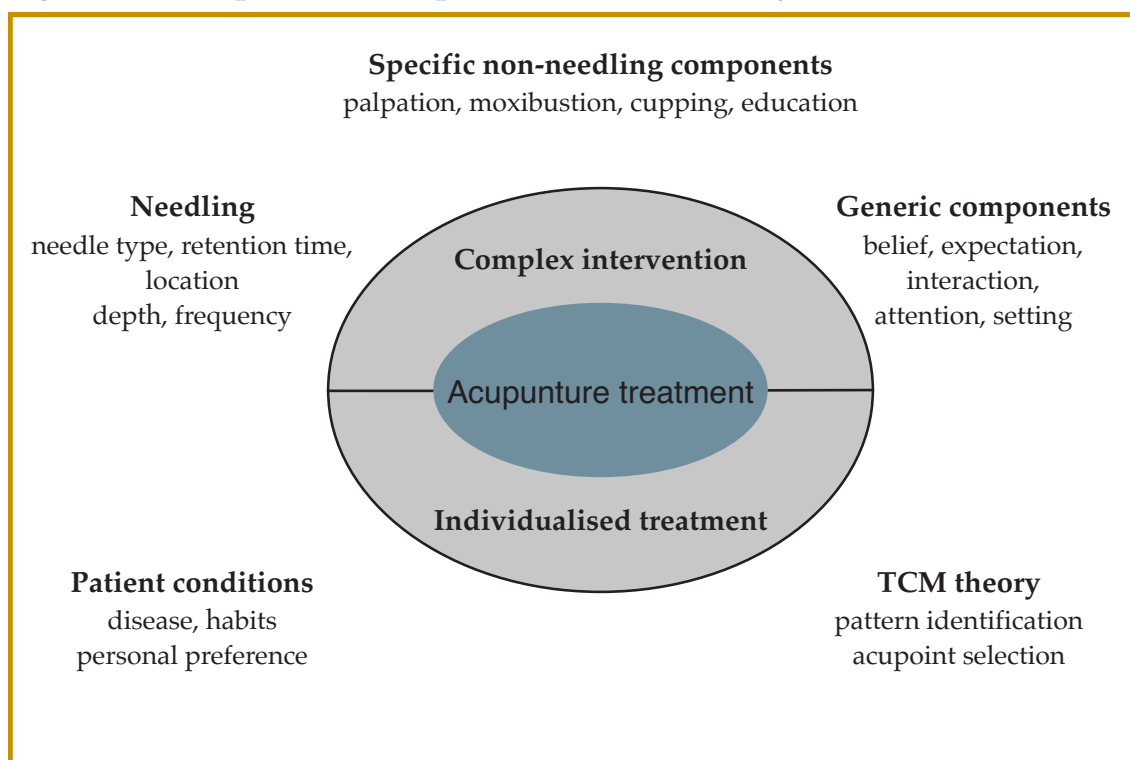
Differences in capacity, resources, interests and requirements of stakeholders will influence research at the local level. Clinical research may also be influenced by the results of mechanism studies and trends within practice. It is unlikely, therefore, that at any point in time acupuncture research will be following a simple linear path, globally. Therefore, the circular conceptualisation is a more appropriate means of synthesising existing knowledge and identifying future avenues for research. Nevertheless, the linear process of establishing comparative effectiveness then investigating component efficacy and mechanism should continue. In addition, comparative effectiveness trials are important as they enable patients and stakeholders to make informed

decisions; and have been identified as a key element needed for research into CAM (Fischer et al., 2014).

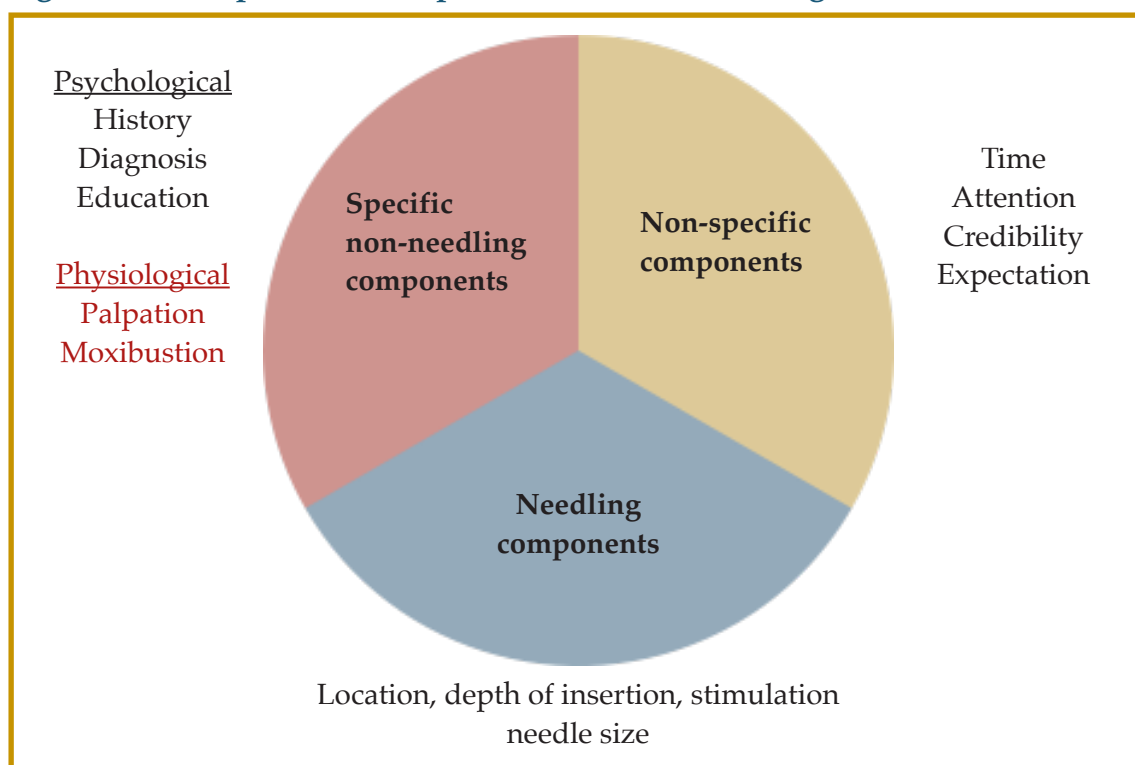
II.3.2. The components of acupuncture - schema

Reproducibility requires an intervention to be described fully, with a clear understanding of its components (Craig et al., 2014). Different schematic illustrations have been proposed to represent the components of acupuncture. For example, one from the China Academy of Chinese Medical Sciences (Luo et al., 2015) another from the SAR (Langevin et al., 2011) ([figure-II.3](#) & [figure-II.4](#)). Both illustrate the complexity of acupuncture. The schema presented by Luo et al (2015) has highlighted TCM theory and pattern differentiation.

Figure-II.3: Acupuncture components China Academy schema (Luo et al 2015)



In the SAR framework notice that the needle components have been placed in a single category, whilst the associated physical techniques, palpation and moxibustion, have been categorised along with specific psychological components ([figure-II-4](#)).

Figure-II.4: Acupuncture components SAR schema Langevin et al (2011)

II.3.3. Accommodation of TBSA clinical reasoning in research

Many studies that claim to have tested TBSA have lacked external validity as they did not incorporate pattern differentiation of that system. This prompted calls for protocols to be developed more rigorously, using a variety of methods including literature reviews and Delphi surveys (Alraek, 2014, Birch, 2004, Birch and Alraek, 2014, Hammerschlag and Stux, 2001, Schnyer and Allen, 2002). There was also a need to improve standards of reporting which led to an extension of the Consolidated Standards of Reporting Trials (CONSORT) (Moher et al., 2010); the Standards for Reporting Interventions in Clinical Trials of Acupuncture (STRICTA) (MacPherson et al., 2010).

There has been an improvement in the quality of acupuncture research, however, a fundamental dilemma remains. Researchers should strive for reproducibility when evaluating a complex intervention (Craig et al., 2014). But, TBSA clinical reasoning is a subjective process ([section-I.3.5.b](#)). Therefore, to gauge the reproducibility of acupuncture treatments it is necessary to understand the reliability of the clinical reasoning process.

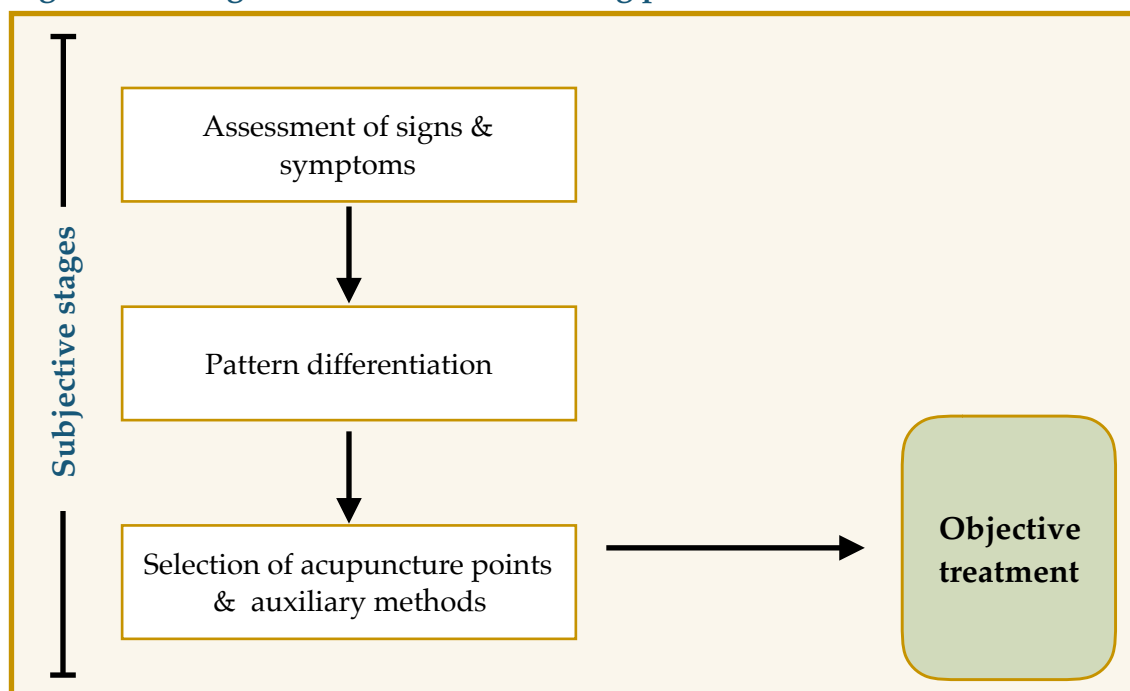
Pattern differentiation is one component of the clinical reasoning process: it is preceded by assessment of the signs and symptoms; and followed by the selection of acupuncture points and auxiliary methods ([figure-II.5](#)). The assessment of signs and symptoms is a subjective process ([section-I.3.5.b](#)). Consequently, there is a need to evaluate practitioners' assessment using inter-rater and intra-rater reliability studies. Reviews found that few studies have been undertaken, and these were relatively small (Bilton and Zaslawski, 2016) (O'Brien and Birch, 2009). Studies have examined two characteristic diagnostic methods of Chinese medicine, the assessment of the tongue and pulse (Bilton and Zaslawski, 2016, Kim et al., 2008a, King et al., 2006, O'Brien et al., 2009c). A few studies have looked at other signs and symptoms (Hua et al., 2012, O'Brien et al., 2009a, O'Brien et al., 2009b). The studies suggest some signs and symptoms are more reliably assessed than others (Hua et al., 2012, O'Brien et al., 2009b). Reliability in assessment of the pulse decreased as the complexity of the task increased (Bilton and Zaslawski, 2016, O'Brien and Birch, 2009). There was some evidence to suggest reliability could be improved after specific training (King et al., 2002).

Studies of the analytical process of pattern differentiation also presented a mixed picture in respect of reliability (Birkeflet et al., 2011, Birkeflet et al., 2014, Grant et al., 2013, O'Brien et al., 2009a, O'Brien et al., 2009b, O'Brien and Birch, 2009). There was some evidence to suggest that training candidates could lead to greater reliability (O'Brien and Birch, 2009). A process of manualisation has been recommended to improve the capacity to standardise and replicate protocols (Schnyer and Allen, 2002).

Once the differential pattern has been identified the final component on the clinical reasoning process is the selection of the treatment methods. These include the selection of acupuncture points, types of needles, use of moxibustion, and auxiliary techniques. Studies have also shown that there was

variability in the selection of acupuncture points (Coeytaux et al., 2006, Hogeboom et al., 2001, Kalauokalani et al., 2001b).

Figure-II.5: Stages of the clinical reasoning process



Literature reviews, practitioner or delphi surveys, followed by practitioner training can help ensure procedures are replicable and have external validity, to a reasonable degree. However, this can only test the views of certain experts or a particular style of practice. Schnyer and Allen stated that “The selection of *optimal* [emphasis added] acupuncture treatment is made on the basis of several decisions that are determined by the training and background of the acupuncturist...” (2002, p626). Describing acupuncture as ‘optimal’ is potentially problematic. At present what constitutes optimal practice is unknown. The uncertainty regarding optimal practice has two aspects.

The first aspect can be conceived as disputes between different styles of acupuncture, or between practitioners, each would claim to be the most effective (section-I.3). The depth of insertion and the necessity to elicit an aching sensation (deqi) are both points of contention within the field of TBSA. Due to a limited appreciation of acupuncture practice, shallow insertion and avoidance of an aching sensation have both been wrongly regarded as placebo (Appleyard

et al., 2014). Practice has evolved alongside wider philosophical and social changes ([section-I.3.2](#)), there is no guarantee that current practice is optimal. Even limiting ourselves to modern practice we are faced with Hume's induction problem (Vickers, 2016). No matter how rigorous the protocol development, or the number of styles tested, there will always be practitioners who believe it does not reflect (their) best practice.

The second aspect is the inherent subjective and skill based nature of acupuncture. The clinical reasoning process is subjective and the delivery of acupuncture a physical skill ([section-I.3.5.c](#)). Therefore, when considering clinical trials we should be aware that a dichotomy of acupuncture/no acupuncture is too simplistic. It is not the same as a drug trial where the active ingredient is or is not present. Delivery of acupuncture should be conceptualised as a distribution around a mean, where optimal acupuncture is seldom, if ever, actually delivered. In health, science must deal with probability rather than clear-cut cause-and-effect relationships (Rosenberg, 2005).

In summary, the clinical reasoning process can be considered as three subjective components - assessment of signs and symptoms, pattern differentiation, selection of points and methods - that lead to a treatment protocol that can be objectively described terms of the physical variables ([figure-II.5](#)). There is potential to increase the reliability of each of these stages through training and manualisation. However, this approach is limited by the lack of a clear understanding of what constitutes optimal acupuncture.

II.4. New research framework - Component analysis

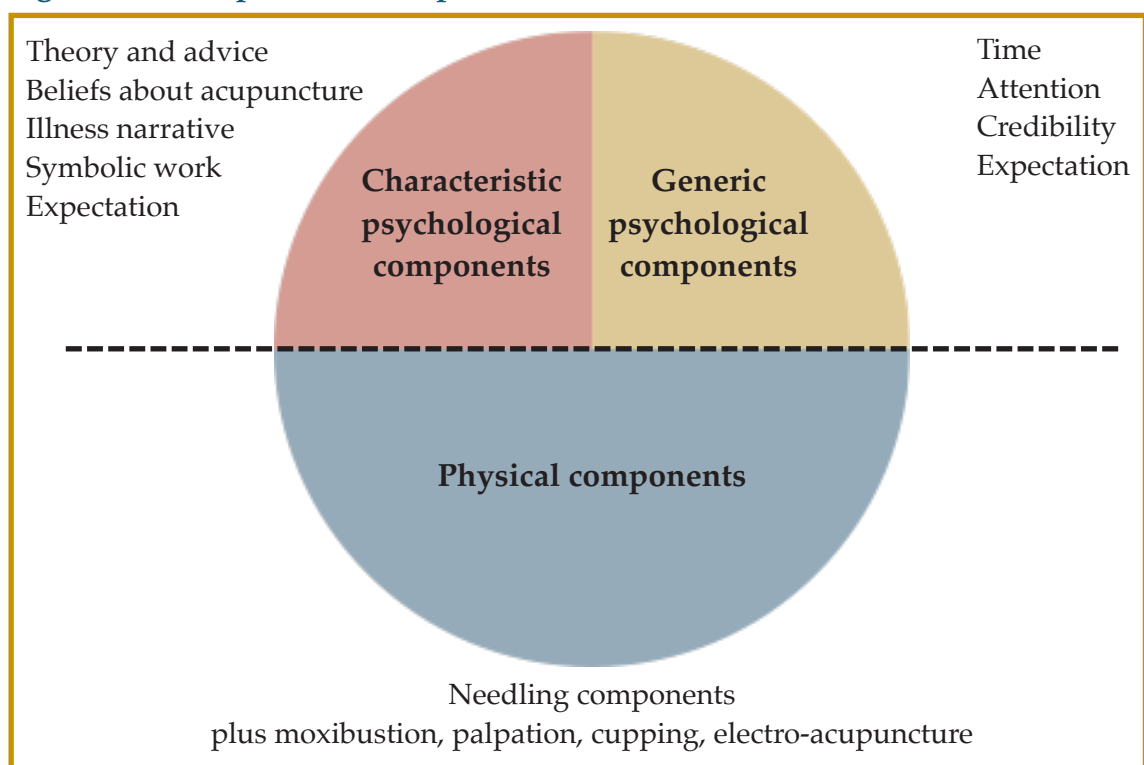
The research framework presented in this section is part of the new knowledge developed for this PhD project. The framework underpins the rationale for the various phases; the expert interviews ([chapter-V](#)), the practitioner survey ([chapter-VI](#)) and pilot study ([chapter-VIII](#)), as well as informing the overview of acupuncture for osteoarthritis of the knee ([chapter-VII](#)). There are three innovations. The categories of acupuncture components in the SAR schema (Langevin et al., 2011) are adjusted. The concepts of micro-meso-macro theories are incorporated to support the development of an overall theory of acupuncture. The third innovation uses component analysis to understand the clinical reasoning process. This provides a method to accommodate the tension between external validity and replicability. It also offers a means by which sham acupuncture procedures can be assessed ([chapter-VII](#)).

II.4.1. New schema of the components of acupuncture

The new schema is an adaptation of the SAR schema ([figure-II.4](#)). There are three categories, physical components plus characteristic and generic psychological components ([figure-II.6](#)). There are two principal reasons for adapting the SAR schema in this way. Firstly, to reflect the key question of whether the effects are mediated psychologically or via the physical processes ([section-II.2.2](#)). This central issue has been clarified by categorising the components as either physical or psychological. Secondly, the separation of needling from moxibustion was not soundly based on TBSA ([section-I.3.4](#)). A schema that categorises moxibustion along with needling components is more reflective of practice, according to TBSA. Thus, not only acupuncture needling components but also moxibustion, palpation, cupping, guasha and electro acupuncture, have been included in the physical components category. The terms placebo and specificity are highly problematic ([section-I.4](#)). This has obscured the fact that some psychologically mediated effects are specific characteristics of acupuncture ([section-I.4.2.d](#)). Therefore, the spectrum of

psychological factors, often collectively conceived as the placebo response, have been divided into characteristic and generic psychological factors. Generic psychological components are those that are shared with other therapeutic interventions, such as time and attention. Characteristic psychological component are those that relate specifically to acupuncture; and are shared with few, if any, other interventions: for example, beliefs regarding acupuncture, the symbolic experience of having acupuncture, and illness narratives in terms of TEAM. Some components may have characteristic and generic aspects, for example the expectation of trying a new therapy and the expectation specifically of acupuncture.

Figure-II.6: Acupuncture components - new schema



Advice based on TEAM may influence attitudes towards acupuncture treatment as a whole, and therefore can be considered a psychological component. Advice, of course, may also lead to behavioural change in terms of diet and exercise, which are physical. Behavioural change is not included in [figure-II.6](#). The new schema focusses on components of acupuncture present during treatment sessions.

Variables and constructs

[Figure-II.7](#) expands the schema to illustrate the components of acupuncture in more detail. The physical components are described as *variables*. The term *construct* is used by psychologists instead of *variable*, as construct has a more abstract connotation, rather than being a specifically defined term (Creswell, 2008). Therefore the components that may have *psychologically mediated effects* have been described as *constructs*.

The physical components of acupuncture - the variables, are presented on the right hand side of the diagram. Measuring the variables should be relatively straightforward in comparison to measuring the constructs. The constructs need to be operationalised through the development of a range of instruments, that can be utilised to investigate their effects. Experiments that test the physical variables of acupuncture, and successfully blind participants, can demonstrate non-psychological effects. Arguably given the value currently placed on physical variables ([section-I.4.2.e](#)), evidence that suggests a non-psychological basis will be especially influential. Consequently, there is a critical divide between the variables and constructs. The new schema is designed to illustrate this issue. The protocol tested in the pilot study ([chapter-VII](#)) investigates a physical variable.

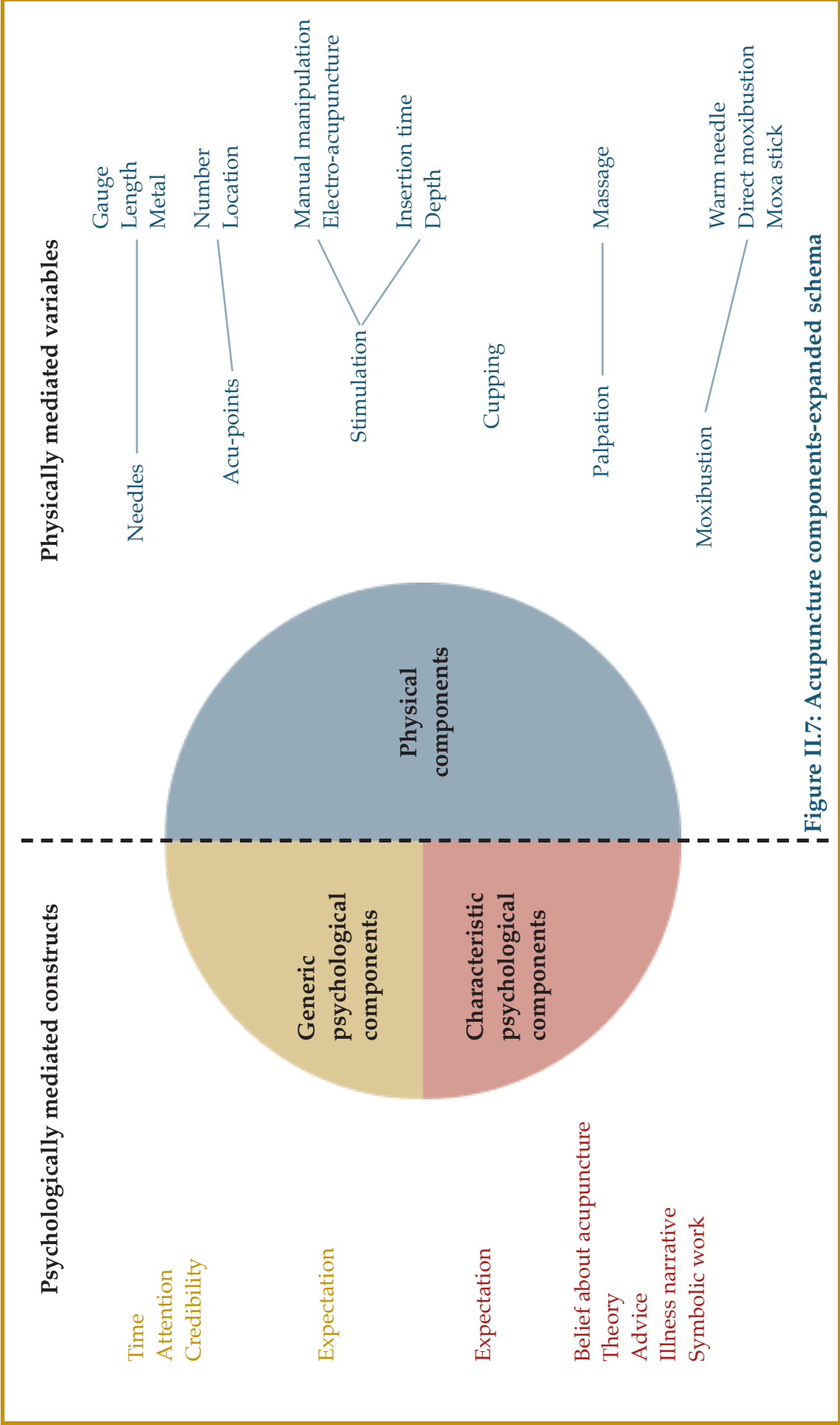


Figure II.7: Acupuncture components-expanded schema

II.4.2. Micro - meso - macro theories in the new framework

TBSA does not have a falsifiable theory that can be subject to scientific investigation ([section-II.2.1.b](#)). The gold standard placebo controlled trial is not viable due to the lack of a placebo acupuncture and the diversity of practice ([chapter-I](#)). Ideally, a scientific experiment investigates the effects of a single independent variable on a dependent variable, to establish causality. Research can proceed by identifying components which can feasibly be tested by experimental methods. These are *component efficacy trials*, rather than efficacy trials of acupuncture in its entirety.

Theories vary in their breath of coverage and can be categorised into different levels: micro-level theories provide explanations limited to small slices of time, space or numbers of people; macro-level theories explain larger aggregates, such as social institutions, cultural systems and whole societies; meso-level theories link the micro to the macro-level (Neuman, 2009). This understanding of theories has been derived from the social sciences, however, it is useful to adopt the notion for acupuncture research.

Knowledge of acupuncture can be used to identify situations where ‘if-then’ statements apply. These can be considered *micro-level theories*. The micro-level theories can be tested, and potentially falsified ([section-II.2.1.b](#)), by means of experiment. In a clinical setting these experiments are *component efficacy trials*. Ideally the experimental group will differ from the control group in respect to a single construct or variable that maybe present/absent, or have different parameters or quality. Related situations can be identified that generate similar micro-level theories which, if verified, can then be used to develop *meso-level theories*.

As an example, consider the following: There is evidence that P6 *neiguan* 内关 is effective in reducing post operative nausea and vomiting (PONV) (Cheong et al., 2013, Lee et al., 2015). This is an example of micro-level theory that has been tested. TBSA was used to identify the if-then statement: if P6 is stimulated, then

PONV will be reduced. Other acupuncture points are also thought to be beneficial for nausea and vomiting, or example SP4 *gongsun* 公孙 (Deadman et al., 2007). However, none of these have been investigated to the same extent as P6 (Streitberger et al., 2006). Basic science and clinical trials can both be used to try and establish whether SP4 is also effective for nausea and vomiting, and whether there is any connection between the P6 and SP4. This would establish a meso-level theory. The meso-level theories can then be used to try to construct a *macro-level theory* of acupuncture.

II.4.3. Key principle - independence of components

It is possible for some parts of Chinese medicine to be true whilst others are false. On a practical level, this means that some components of acupuncture may be clinically effective whilst others are not. Practitioners already approach TBSA in a manner that typically rejects some traditional ideas, whilst accepting others. They tend to sit somewhere on a spectrum between wholesale acceptance and complete rejection. Even though Western medical acupuncture (WMA) does not include concepts such as Yin-Yang, it does incorporate acupuncture points drawn from TBSA (White et al., 2008, White and Editorial-Board, 2009). Practitioners who incorporate Yin-Yang within their praxis do not necessarily subscribe to the astrological theory Stems and Branches (*tiangan dizhi* 天干地支), that is also part of Chinese medicine. Somewhat ironically, a treatise on folk medicine written in 1759 did not include discussions on Yin-Yang (Unschuld, 2010). This indicates that, for the lower echelons of Chinese society, medical practice was not inextricably linked to Yin-Yang theory. So western medical acupuncturists are not the first to practice without reference to Yin-Yang theory.

Practitioners adopt or reject different components of TBSA based on their own personal beliefs. These personal beliefs may be based on a number of factors such as training, experience or understanding of the evidence base. In research, the evidence relating to specific components must be treated as being independent. This allows the possibility to reject some components whilst

accepting others, as practitioners currently do. Because the components should be treated independently, it implies that evidence regarding one component cannot demonstrate the validity of another component ([section-II.2.1.b](#)).

Applying these principles to the current evidence base, we can again consider the evidence that P6 *neiguan* 内关 is effective in reducing post operative nausea and vomiting (PONV) (Cheong et al., 2013, Lee et al., 2015). Scientific circumspection dictates that this evidence does not demonstrate that acupuncture can be used for nausea due to another cause, such as motion sickness. It also does not provide any evidence that other acupuncture points can be used for different conditions. It would be incorrect to suggest that the evidence for P6 demonstrates that ST36 *zusanli* 足三里 can be used for digestive problems. But the reverse is also true, if after rigorous investigation there is no evidence that ST36 can be used for digestive problems, this does not invalidate the evidence for P6 and PONV. Similarly when some commentators claim that the results for the so-called real acupuncture and sham acupuncture implies 'the entire philosophy of acupuncture falls apart' (Singh and Ernst, 2008), they are overstating what can be deduced from the evidence. Their claim rests on a questionable assessment of sham procedures (Appleyard et al., 2014). However, even if this were not the case, clinical evidence for chronic pain does not invalidate the evidence for P6 and PONV. We may discover that point specificity is less relevant for chronic pain than it is for nausea and vomiting.

II.4.4. Component efficacy - Risks and challenges

[Figure-II.7](#) illustrates a considerable number of components that may potentially have a bearing on the overall effect. The efficacy of any single component should be smaller than the overall effectiveness. Hence, effect sizes within component efficacy trials are likely to be relatively small and there is a risk of a type II error. Small effect sizes also have implications for assessment of clinical significance and the practicalities of component efficacy.

Clinical significance in component efficacy

Small effect sizes, that are statistically significant, are not necessarily clinically significant (Tubach et al., 2012). Typically 0.2 is considered a small effect (Field, 2009). The differences between ‘sham’ acupuncture and ‘real’ acupuncture were 0.23 (95% CI, 0.13-0.33), 0.16 (95% CI, 0.07-0.25) and 0.15 (95% CI, 0.07-0.24) SDs for back and neck pain, osteoarthritis, and chronic headache, respectively (Vickers et al., 2012). These have been described as clinically insignificant (Colquhoun and Novella, 2013). It has been suggested the single end-point outcome measures, that imply a linear cause and effect relationship, may not be appropriate for complex interventions (Paterson et al., 2009). In component efficacy an if-then clause can be tested, with the objective to understand a cause and effect relationship. Therefore single endpoints outcome measures are appropriate, but the degree of change considered clinically significant must be re-evaluated.

Despite frequently being described as acupuncture versus placebo, many studies have actually been component efficacy trials (Fønnebø et al., 2007). This is because they compared the effects of using different needling parameters, depth of insertion and location ([section-I.4.4](#)). They did not use ‘placebo’ procedures (Appleyard et al., 2014). Evidence that point location influences clinical outcomes is useful information both for clinicians and scientists researching mechanisms (Lewith and Vincent, 1995). From a scientific point of view, if a different depth of insertion leads to a greater reduction in pain then it begs the question as to why that should be the case. In practice acupuncturists should try to deliver the best possible treatment.

Practical considerations of component efficacy

A lack of adequately powered studies has been misinterpreted as evidence that acupuncture is ineffective. For example, it has been stated: “the bad news for acupuncturists. The Cochrane reviews suggest that there is no significant evidence to show that acupuncture is an effective treatment for any of the following conditions: ...glaucoma.....vascular dementia” (Singh and Ernst,

2008, p77). For vascular dementia the Cochrane review reported that no trials met the standard required to be included in the review (Peng et al., 2007). The review of glaucoma only included one trial, that trial used auricular acupuncture (Law Simon and Li, 2013). This type of analysis maybe an example of interpretative bias towards acupuncture. High-quality evidence is an important precondition before the effectiveness of any therapy can be judged (Altman and Bland, 1995). Components efficacy studies will need to be especially explicit in describing the hypothesis being tested, to avoid the results being misconstrued.

Securing adequate resources to conduct component efficacy trials, that are sufficiently powered, will no doubt be enormously challenging. Therefore, careful consideration of which components to test will be of paramount importance. A translational research approach should be employed to guide this process, utilising not only the knowledge of practitioners but also basic science research.

II.4.5. Accommodation of clinical reasoning in the new framework

To develop the evidence base, acupuncture protocols ought to be replicable (section-II.3). However, acupuncture according to the precepts of TBSA is incompatible with replication (section-I.3.5). In addition, there is a lack of a clear understanding of what constitutes optimal acupuncture (section-I.3.3). This section proposes that component analysis can applied to the clinical reasoning process to develop methods of accommodating these seemingly intractable problems. There are three axioms that underpin this approach box-II.1.

Box-II.1: Axioms of the new research framework

- Optimal practice is currently unknown, therefore, research should investigate the role of specific components to move towards an understanding of optimal practice
- External validity will inevitably be compromised in controlled trials that use fixed or semi-standardised protocols. An understanding of which components have been omitted or restricted is a pre-requisite for effective evaluation of how these trials reflect clinical practice.
- Sham acupuncture procedures will inevitably retain some components of acupuncture. An understanding of which components have been retained is a pre-requisite for effective evaluation of these procedures.

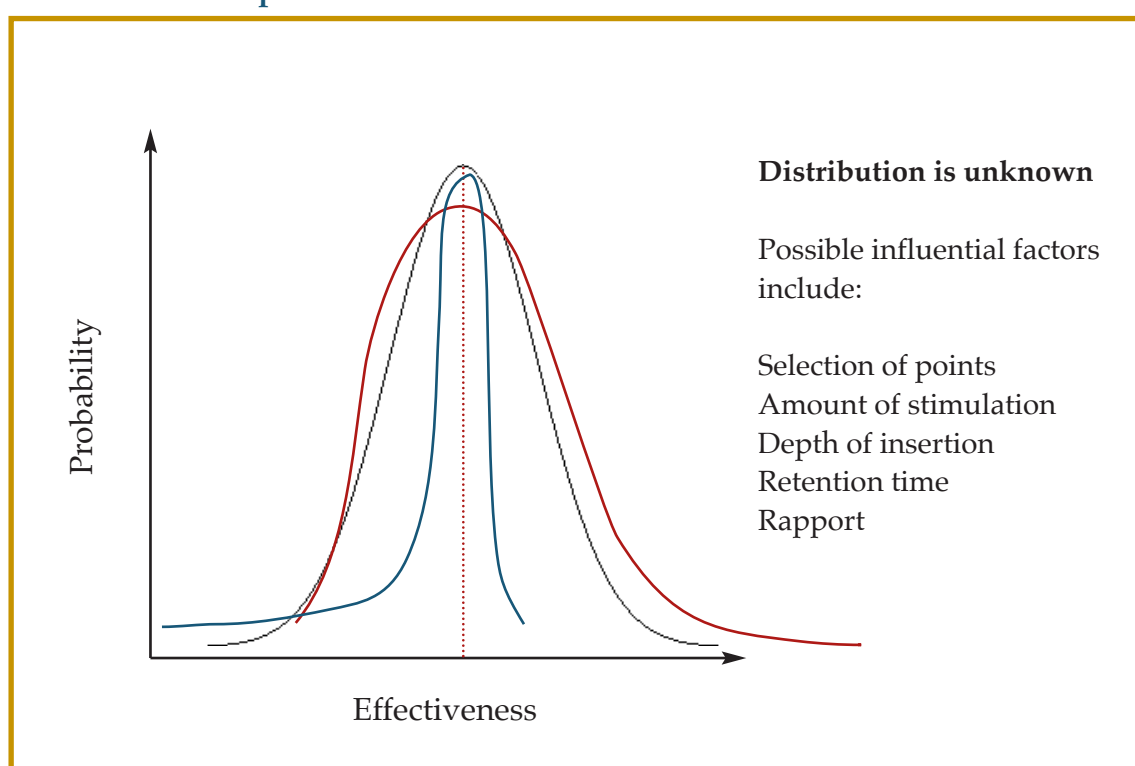
II.4.5.a. Optimal acupuncture

Within clinical research, acupuncture should not be considered as a simple dichotomy of being delivered or not ([section-II.3.3](#)). Effectiveness is a distribution that is dependent upon how well a number of components are delivered. Components such as selection of appropriate points, accurately locating points, employing the correct amounts of stimulation, and a variety of other factors such as the therapeutic relationship, all play a role. The effectiveness may also be dependent on the style of practice and the individual skill of the practitioner ([section-I.3.5.c](#)). In a similar fashion to sports people or musicians, an individual practitioner will not be able to replicate their best performance on every occasion. Theoretically, how effective any given acupuncture treatment is for an individual patient will fall somewhere on a probability distribution. Optimal acupuncture might only occur on rare occasions when the practitioner delivers the treatment exceptionally well, leading to a marked improvement compared to most treatments. Alternatively most treatments may deliver more or less the same result. At present the shape of this distribution is unknown ([figure-II.8](#)).

In comparative effectiveness research (CER) this is less problematic as stakeholders wish to know the effectiveness of the current acupuncture workforce. If there are highly skilled practitioners elsewhere, this does not help

patients. For understanding best practice and mechanisms, these are issues that must be considered. Best practice can be investigated by allowing one component of the intervention to vary whilst balancing the other components between arms of a trial (Cardini et al., 2006, Craig et al., 2014, Fønnebø et al., 2007). In other words, by undertaking component efficacy studies. These scientific methods can be applied to TBSA to resolve disputes that have existed for centuries.

Figure-II.8: Three hypothetical probability distributions for the effectiveness of individual acupuncture treatments



II.4.5.b. Pattern differentiation - A staged process

Figure-II.5 illustrates three components of clinical reasoning: assessment of signs and symptoms, pattern differentiation, selection of points and methods of stimulation. The pattern differentiation process can be further divided into stages. Acupuncture textbooks contain examples of differential patterns with associated signs and symptoms (Maciocia, 1994, Xu et al., 1988). For each pattern there are suggested acupuncture points and auxiliary methods. In

practice, acupuncturists will adjust these standard treatment protocols depending on the individual clinical picture of the patient (section-I.3.5). Consequently, pattern differentiation and the subsequent implementation of acupuncture can be viewed as a process of refinement. In theory, as one moves from a crude treatment protocol to an increasingly individualised and refined intervention the effectiveness should increase. In this framework the process of refinement has been split into four stages. Figure-II.9 illustrates these four stages and has used a normal distribution curve for simplicity, however, the actual distributions may be different (section-II.4.5.a). As one moves from Stage 1 to Stage 4 the external validity increases, but the replicability of the protocol decreases.

Stage 1: The protocol is based on acupuncture points thought to be effective for a particular bing-illness. The use of P6 for nausea and vomiting is an example on this kind of protocol.

Stage 2: Pattern differentiation is incorporated into the protocol. The protocol is flexible but with some controls to enable replicability and standardisation. These controls may include a pre-defined lists of possible patterns, with corresponding acupuncture points and other methods. Practitioners delivering the acupuncture should receive training.

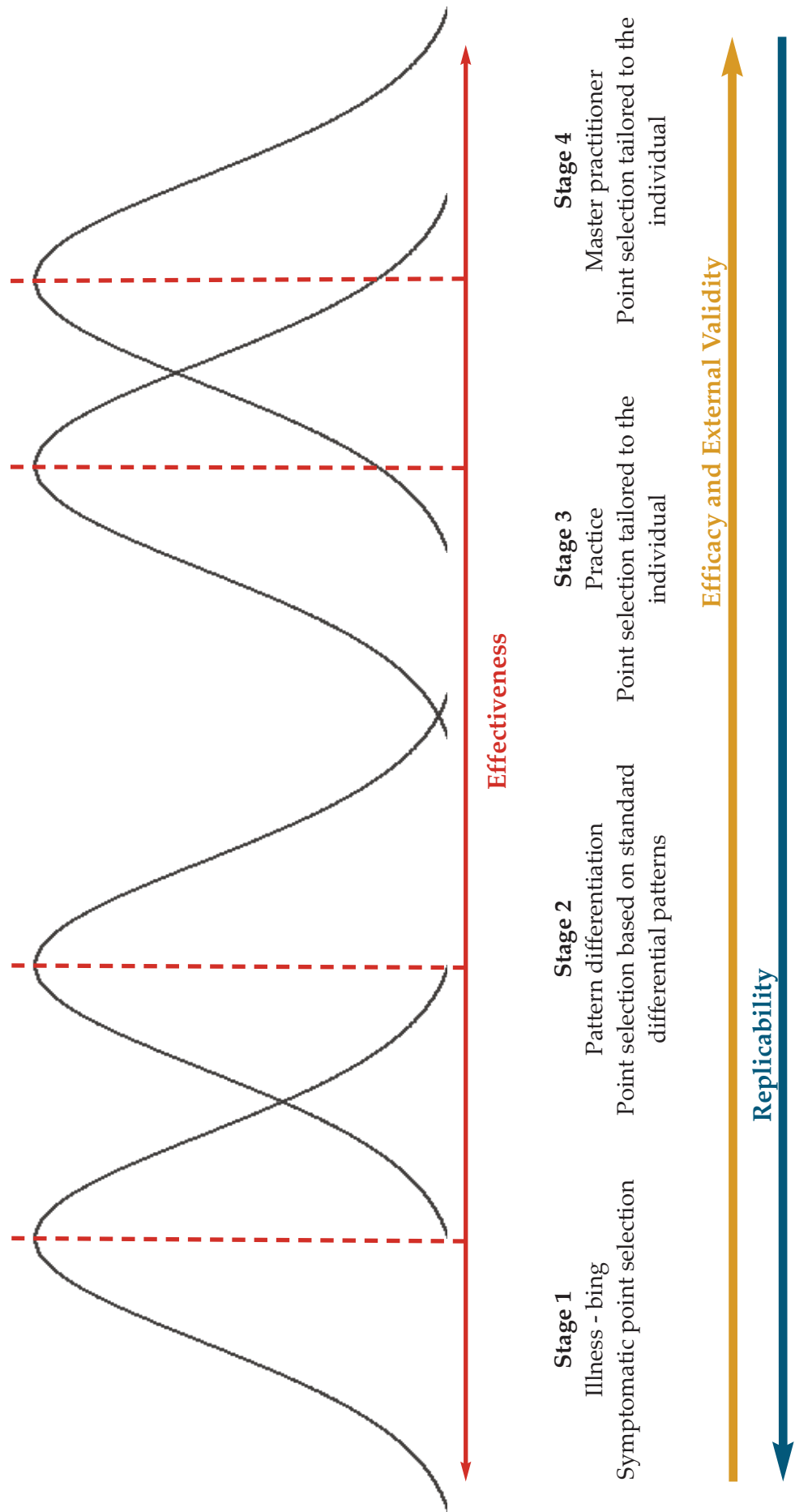
Stage 3: A qualified practitioner is allowed to deliver acupuncture as they would in normal practice

Stage 4: A master practitioner is allowed to deliver acupuncture as they would in normal practice

The number of associated differential patterns varies between bing-illness. A standard Chinese medicine textbook outlines 20 different patterns for tiredness, but only 4 patterns for sinusitis. Furthermore, the four patterns for sinusitis share a number of recommended points in common (Maciocia, 1994). This suggests that the potential heterogeneity, in point selection and stimulation

methods, at stage 3 is relatively small for sinusitis when compared to tiredness. A semi-flexible protocol for a binge-illness with fewer associated patterns should have a greater degree of external validity. In other words, there will only be minimal differences between stage 2 and stage 3. It should also be remembered that binge-illness do not necessarily have a direct relationship with a biomedical disease category ([section-I.3.5.a](#)). For example, a cohort of individuals with diabetes type II would not necessarily be considered to have the same binge-illness in TBSA. If participants have different binge-illness, then there maybe more differential patterns to consider, increasing the heterogeneity.

Figure II-9: Pattern differentiation - Stages



Two other factors may influence replicability, inter-rater reliability of assessing signs and symptoms and the relative importance of the ben-root. The process of assessing the signs and symptoms is subjective (Alraek, 2014) (section-I.3.5). However, certain signs and symptoms have a greater inter-rater reliability than others (section-II.3.3). For example, research suggests there is substantial agreement regarding the speed of the pulse but only slight agreement for the colour around the eyes (O'Brien et al., 2009c). Conditions where the pattern differentiation revolves around key signs and symptoms that can be assessed reliably will be more amenable to replication. It may also be the case that not all elements of a pattern differentiation are equally important. It has been suggested that a protocol for lower back pain would reasonably reflect practice because TCM practitioners did not view the more complex, and less predictable, root pattern differentiation to be important (Sherman et al., 2001b). Because of these characteristics certain conditions may provide a better opportunity to answer specific questions. If the objective is to investigate the effects of a variable, such as depth of insertion, then conditions that have minimal differences between stage 2 to stage 3 may be more suitable. As a semi-flexible protocol would have a high degree of external validity. It is important to retain as much external validity as possible otherwise the acupuncture maybe too ineffective to demonstrate any change. The attributes that should lead to minimal differences between stage one and three have been summarised in box-II.2.

Box-II.2: Attributes of bing-illnesses that maximise replicability

- ▶ Few associated differential patterns
- ▶ Similar points are recommended for the associated patterns
- ▶ Key signs and symptoms have high inter-rater reliability
- ▶ Complex components of the clinical reasoning process such as the ben-root are not considered important

If one wished to investigate the pattern differentiation process, then a trial should compare a stage 1 to a stage 3 protocol. Under these circumstances it may be better to select a relatively complex condition. Karner et al (2013) compared individualised classical acupuncture to a modern semi standardised acupuncture; these could be considered stage 3 and stage 2 protocols, respectively. The data suggested that there is a specific effect of using a stage 3 protocol for knee mobility.

II.4.6. Translational Research

A translational approach has been widely recommended, with three domains informing and being informed by each other; mechanistic studies, clinical studies and clinical practice (Langevin et al., 2011, Witt et al., 2012). In this framework TBSA is also considered an active research domain. The rationale for utilising TBSA is set out below. In the subsequent section it is argued that a translational research approach is essential for the proper evaluation of clinical research ([section-II.4.6.b](#))

II.4.6.a. Rationale for utilising TBSA

Attitudes towards the role of TBSA in the research and practice of acupuncture are, at times, diametrically opposed. A better understanding and greater utilisation of TBSA has been advocated (Cardini et al., 2006, Langevin et al., 2011, Luo et al., 2015, Witt, 2011). Yet, at the same time, some practitioners of acupuncture have already consciously rejected the traditional theories (White and Editorial-Board, 2009). Whilst other practitioners feel that traditional TBSA cannot be investigated using clinical trials (Quah, 2003). This methodology takes as a priori that a comprehensive strategy to investigate acupuncture should utilise an in-depth understanding of TBSA.

The rationale for including TBSA as a priori within the research framework is, simply, that some components of acupuncture practice already have an evidence base. This being the case, not only would it be premature to reject other components without proper investigation, but it would also be a missed

opportunity for science. For simplicity only the evidence P6 for PONV and chronic pain is discussed, rather than a comprehensive assessment of acupuncture for all conditions.

The meta-analyses with greatest statistical precision to date shows acupuncture is more effective than conventional care for chronic pain conditions (Vickers et al., 2012, Vickers et al., 2017). The extent to which these effects are due to characteristic psychological components or the physical variables is yet to be fully understood ([section-I.4.2.d](#)). Nevertheless, the notion of inserting needles to reduce pain has come from TBSA. Moreover ‘real’ acupuncture outperformed ‘sham’ acupuncture (Vickers et al., 2017). Participants were blinded to group allocation, so the difference was due to the physical variables of location and depth of insertion of the needles ([section-I.4.4](#)). The locations and depth of insertion of the ‘real’ acupuncture was guided by TBSA.

TBSA was also essential for the production of the knowledge that P6 can be used for PONV (Cheong et al., 2013, Lee et al., 2015). This is because, even now, there is no clear understanding of the biomedical mechanism that explains how the stimulation of P6 leads to a reduction in nausea and vomiting (Streitberger et al., 2006). The evidence for P6 is based on trials where participants were also blinded (Cheong et al., 2013, Lee et al., 2015). Thereby, providing evidence that the component of acupuncture P6 has a physical basis for its ability to reduce PONV. Without reference to TBSA the trials investigating P6 simply could not have been envisaged.

A question that arises from the P6 research is: do other acupuncture points have specific effects? At present other acupuncture points have not been investigated to the same extent. This is a problem of paucity of evidence, not negative results. Is it purely serendipity that the first point to be extensively investigated is also the only acupuncture point that has a specific function? Imaging techniques have been used to investigate correlations between the areas of the brain affected by needling specific acupuncture points and the TBSA functions

ascribed to those points (Huang et al., 2012). While some correlation was identified, it was not 100%. Assuming that TBSA contains errors, there is the possibility of some 'noise' from functions ascribed to acupuncture points that are invalid. P6 is not unique, other acupuncture points also appear to modulate specific areas of the brain (Huang et al., 2012). Common sense would suggest that it is unlikely that P6 is the only location on the body that has a specific effect when stimulated with a needle. P6 provides science with an opportunity to investigate a phenomenon outside its current understanding. There is the potential to investigate the functions associated with other acupuncture points, which may prove useful in furthering the understanding of physiology.

In summary, the notion of inserting needles to reduce pain comes from TBSA. There is evidence that TBSA ideas regarding location and depth of needle insertion lead to greater clinical benefit are correct (MacPherson et al., 2014, Vickers et al., 2012, Vickers et al., 2017). There is both clinical and basic science research that suggest acupuncture points have specific functions (Cheong et al., 2013, Huang et al., 2012). There have already been unanticipated insights for biomedicine from acupuncture research (MacPherson et al., 2016). Therefore, TBSA should be one of the domains of acupuncture research.

TBSA is in a state of flux, under the influence of a wide range of factors ([section-I.3.3.b](#)) (Scheid, 2002, Unschuld, 2009). Scholarship and better translations of historical texts may change the way acupuncture is practiced ([section-II.2.1.c](#)). As new knowledge emerges from biomedicine, these may alter perceptions of TBSA concepts. Modern practice may eschew certain components of TBSA theory, even among TBSA practitioners. Conversely, modern practice may adopt new methods based on biomedicine. An overview of acupuncture for any condition should encompass an understanding of the ways in which practice and theory have evolved. Comparing and contrasting theory with practice could highlight areas for future research, which may shed light on the clinical reasoning of acupuncturists.

II.4.6.b. Explanation of the need for a translational approach

Within this framework a translational approach is the synthesis of four domains: clinical and mechanism research, practice and TEAM. There is no simple chronological relationship between these four domains. Each domain can, and should, be influenced by the other three.

Knowledge from the domains of theory and practice can support the evaluation of clinical research, help shape the research agenda, as well as provide insights into clinical reasoning. This knowledge is necessary to understand the components of acupuncture for any given condition. One of the axioms of this research framework is that fixed or semi-standardised protocol will compromise external validity ([section-II.4.5.b](#)). Clinical trials can be evaluated with an appreciation of which components, used in practice, have been omitted or restricted. Similarly, sham procedures always retain some elements of real acupuncture. Knowledge of the components and their relative importance for the bing-illness under investigation can be used to estimate the risk of bias of sham acupuncture ([box-II.2](#)).

External validity is not the same as generalisability, even for clinical research where there are no restrictions on the acupuncture delivered. Generalisability can only be properly assessed with an awareness of the homogeneity-heterogeneity of practice. If for example, a protocol is developed to test acupuncture using style X, this will need to be complemented with research that evaluates the homogeneity of practice for style X practitioners. Naturally, if it is shown that style X is heterogeneous in practice, then there will be limitations on the generalisability of the results. Conversely, if it is shown that in practice there are only minimal differences between style X and style Y, in terms of their physical components, then a trial is potentially generalisable to both styles. Consequently, both efficacy and effectiveness clinical research can only be properly evaluated within the context of an overview that includes knowledge of the theory and practice.

There are multiple reasons why mechanism studies are important for acupuncture research. There are risks and challenges associated with component efficacy ([section-II.4.4](#)), basic science can support and guide the clinical research agenda. Mechanism research can be readily be linked to component efficacy studies (Lewith and Vincent, 1995), and potentially provide insights for biomedicine (MacPherson et al., 2016). Evidence of mechanisms may be especially influential in the wider debates regarding science and pseudoscience that affect attitudes towards acupuncture ([section-II.2.1.b](#)). And finally, mechanism studies may also help illuminate optimal practice.

TEAM, in general, and TBSA, in particular, are not static bodies of ancient knowledge. They have already been influenced by biomedicine. There are ongoing internal debates regarding best practice, these can potentially be resolved by clinical and mechanism research. Theories and key concepts are being re-evaluated in the light of historical and linguistic research, which in turn informs the other three domains.

Box-II.3: Summary of the attributes of the new research framework**In line with previous recommendations the framework advocates:**

- ▶ A non-hierarchical circular relationship between methods as the most suitable approach for acupuncture.
- ▶ A translational research approach
- ▶ Rigorous methods of protocol development to increase external validity
- ▶ The importance of describing acupuncture clearly in terms of its components

Novel attributes of the research framework are:

- ▶ A new schematic representation of the components of acupuncture, which clearly separates physically mediated from psychologically mediated components.
- ▶ Emphasises TEAM theory as an active research domain
- ▶ The introduction of the concept of micro, meso and macro theories with the principles that:
 - Experimental methods can be used to test micro theories
 - These can be combined to form meso theories, eventually an overall macro theory of acupuncture
 - Micro theories, related to specific components, must be conceived as being independent
- ▶ The conceptualisation of pattern differentiation as a staged process - which facilitates the design and interpretation of research.

Axioms

- ▶ *There is no single theory of TEAM that can be falsified by means of scientific investigation.*
- ▶ *Elements of TEAM are independent, some maybe true whilst others false*
- ▶ *Optimal practice is currently unknown, therefore, research should investigate the role of specific components to move towards an understanding of optimal practice*
- ▶ *External validity will inevitably be compromised in controlled trials that use fixed or semi-flexible protocols. An understanding of which components have been omitted or restricted is a pre-requisite for effective evaluation of these trials.*
- ▶ *Sham acupuncture procedures will inevitably retain some components of acupuncture. An understanding of which components have been retained is a pre-requisite for effective evaluation of these procedures.*

II.5. The Project overview

This section provides an overview of the methods used, and explains the rationale of the various elements of the project ([figure-II.10](#)). The project can be viewed as having four phases. Phase 1 - understanding acupuncture research: Phase 2 - literature reviews: Phase 3 - acupuncture in practice: Phase 4 - completion of aims: translational analysis and pilot study.

The initial inspiration for the project was the desire to test whether the additional use of moxibustion would be more effective than needle acupuncture alone. Evidence that acupuncture effects are mediated by physical rather than psychological components has important implications beyond the confines of an individual study ([section-II.2.2](#)). This project is primarily concerned with the physical variables of acupuncture.

II.5.1. Phase 1: Acupuncture research (chapters-I/II)

Phase 1 consists of the first two chapters. [Chapter-I](#) discusses the paradoxes of acupuncture research and two key challenges - incorporating TEAM and placebo controls in clinical trials. The response to these challenges is set out in [chapter-II](#). This includes an evaluation of the relationship between TEAM and science, as well as the development of a new research framework.

In phase 1 the literature on challenges and controversies associated with acupuncture research are reviewed. This was achieved by first examining review articles (Birch, 2004, Birch et al., 2004, Langevin et al., 2006, Langevin et al., 2011, MacPherson et al., 2008, Napadow et al., 2008, Schnyer et al., 2008, Wayne et al., 2009). The reference lists were used to identify relevant material. Subsequent searches were made using Medline, CINAHL complete, AMED, PsycINFO, Google scholar and search terms such as acupuncture and placebo, sham, methodology. Searches were also made for other articles by authors in the field. In addition, PhD supervisors suggested further reading.

II.5.2. Phase 2 Literature review (chapters-III/IV)

The primary objective of phase 2 was to identify a suitable condition for the pilot study. The selection of osteoarthritis of the knee was based on a number of criteria. The first was whether the disease burden of osteoarthritis of knee and the current treatment options warranted investigation of a novel therapy (section-III.2). For complex interventions overall effectiveness should be established first, then attention can turn to fine tuning the intervention (Craig et al., 2014). Thus, whether comparative effectiveness evidence suggested that acupuncture is beneficial for osteoarthritis of the knee was the second criterion. A preliminary assessment of the literature indicated that there was overall effectiveness (Manheimer et al., 2010, Vickers et al., 2012). A review of systematic reviews, presented in chapter-IV, was used to evaluate the evidence-base more thoroughly.

The new research framework proposes that certain conditions are better suited to investigate the physical variables. For such conditions, a semi-flexible acupuncture protocol may retain a high degree of external validity (section-II.4.5.b). Each bing-illness in TEAM has a number of associated differential patterns. The fewer associated patterns then, potentially, the less variety in actual treatment. In a standard TCM textbook, knee pain has five associated patterns (Maciocia, 1994). For osteoarthritis the number of associated patterns is reduced to three, due to the fact that the knee should not have a feeling of heat (American College of Rheumatology, 1986) and does not move from joint to joint. Ideally, there should be high inter-rater agreement in the assessment of salient signs and symptoms (section-II.3.3). One of the key symptoms in the TCM diagnosis of knee pain is the location of the pain (section-III.5). There should be a relatively high agreement regarding location of the pain among practitioners. Based on this information, it was assumed that osteoarthritis of the knee possessed the characteristics necessary for a semi-flexible acupuncture protocol to have external validity (box-II.2). The narrative review (chapter-III)

assesses the degree to which these assumptions held true across a variety of acupuncture styles.

The fourth criterion was to demonstrate a gap in the existing of an evidence regarding the use of warm needle acupuncture for osteoarthritis of knee. A systematic review of RCTs of warm needle versus needle acupuncture was carried out ([chapter-IV](#)).

The second aim of this project was to establish a comprehensive overview of acupuncture for osteoarthritis of the knee. Clinical research for any condition can only be properly evaluated with knowledge of theory and practice ([section-II.4.6.b](#)). The secondary objective of the literature review ([chapter-III](#)) was to gather the necessary information on the theory. In addition, a secondary objective of the reviews in [chapter-IV](#) was to gather information on the acupuncture protocols used in the clinical trials.

II.5.3. Phase 3 Acupuncture in practice (chapters-V/VI)

Phase 3 consists of two closely related pieces of primary research; expert interviews ([chapter-IV](#)) and a practitioner survey ([chapter-IV](#)). The expert interviews were used in conjunction with the review of acupuncture theory ([chapter-III](#)) in the development of the survey questionnaire. The primary objective of both practitioner survey and expert interviews was data collection on acupuncture in practice for osteoarthritis of the knee. This data from the practice domain is an integral part of the translational analysis ([chapter-VII](#)).

II.5.4. Phase 4 Aims: Translational analysis and pilot study

Translational analysis of acupuncture for osteoarthritis of the knee (VII)

The evidence-base for acupuncture for osteoarthritis of the knee is re-evaluated in [chapter-VII](#). Two of the axioms of the research framework are: external validity will inevitably be compromised in controlled trials that use fixed or semi-flexible protocols; and sham acupuncture procedures will inevitably retain some components of acupuncture ([box-II.1](#)). Systematic reviews have not assessed the quality of acupuncture nor have they evaluated the risk of bias from the placebo procedures (Hróbjartsson and Gøtzsche, 2010, Madsen et al., 2009, Manheimer et al., 2010, Manyanga et al., 2014, Vickers et al., 2010). Acupuncture in clinical trials for osteoarthritis of the knee has been evaluated in terms of ‘adequacy’ (Ezzo et al., 2001, Purepong et al., 2012). However, at present there is no definitive method to evaluate the quality of acupuncture in clinical trials (Smith et al., 2011). There is also no means to evaluate the risk of bias from sham procedures (Appleyard et al., 2014)

A novel method of assessing external validity and risk of bias from sham acupuncture is presented in [chapter-VII](#). This assessment is founded on the analysis of acupuncture in terms of its components established in phases two and three of this project. Data drawn from the practitioner survey is used to establish benchmarks and estimate the impact on external validity and risk of bias. This work sheds new light on the existing evidence-base.

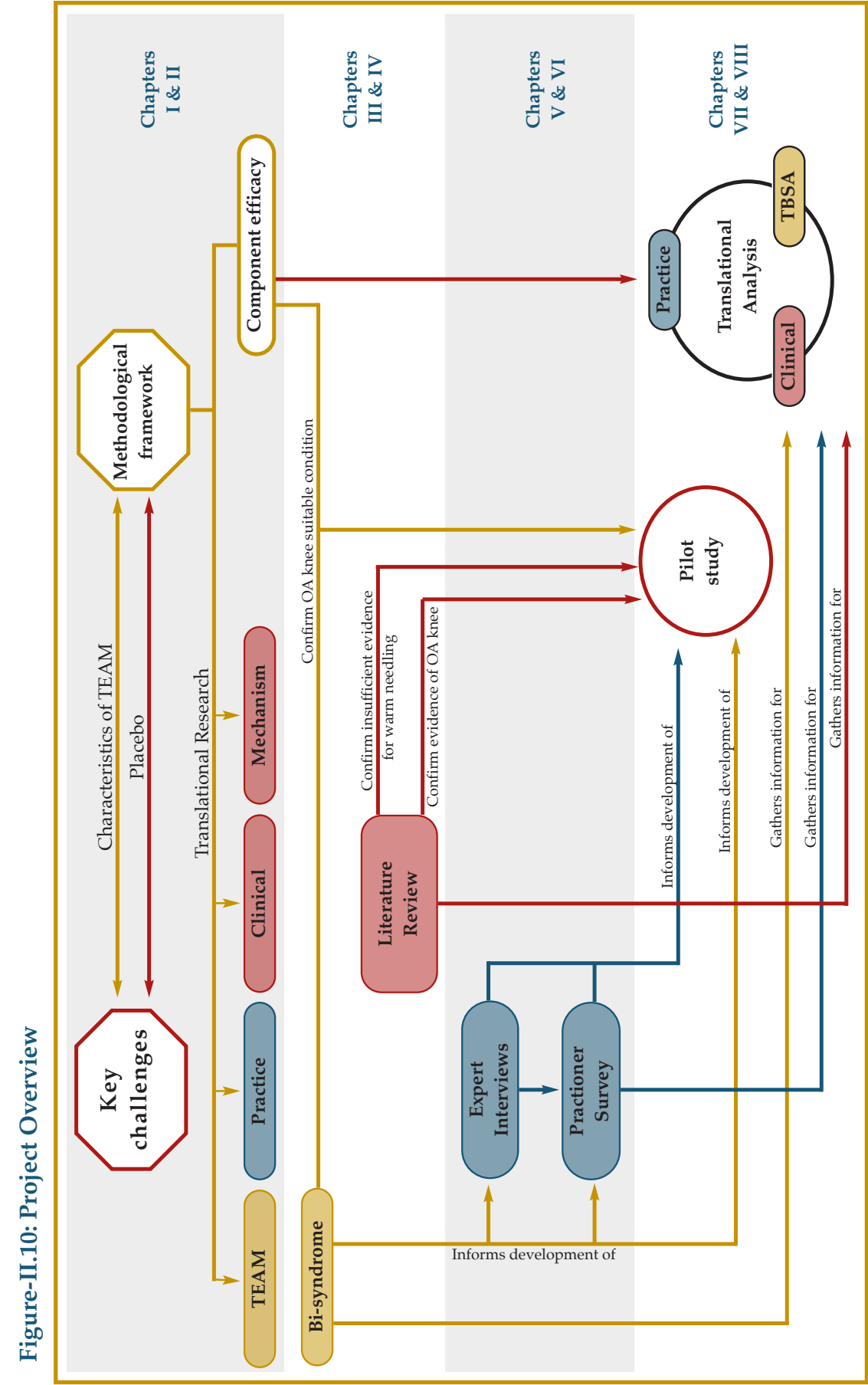
The generalisability of the clinical research is also evaluated. Again, this is supported by the practitioner survey data, and gives an indication on the homogeneity and heterogeneity of practice ([section-I.3.3](#)). In addition, the domains of theory and practice are compared and contrasted, using information derived from phases two and three of this project. Possible trends in the way acupuncture practice is developing are identified.

Pilot study (VIII)

The pilot study was a randomised, controlled, parallel-group 2-armed clinical trial. It was designed so that the outcome administrator, participants and primary acupuncturist would be blinded to patient allocation. Participants were randomised into two groups, they received either warm needle acupuncture or needle acupuncture. As a pilot study the key objective was to test all aspects of the protocol with the ultimate aim of conducting an RCT to test the hypothesis in box-II.4.

Box-II.4: Hypothesis of prospective RCT

Warm needle acupuncture leads to greater clinical benefits for those with osteoarthritis of the knee than the use of needle acupuncture alone



Box-II.5: Objectives of phases 2 and 3**Chapter-III: Narrative review****Preliminary review**

- Assessment of the disease burden and existing therapeutic options

Main review

- Evaluate whether osteoarthritis of the knee met the criteria for the pilot study
- Gather information on theory for the translational analysis of acupuncture for osteoarthritis of the knee (chapter-VII)
- Preparation for the expert interviews (chapter-V)
- Support the development of the practitioner survey (chapter-VI)

Chapter-IV: Systematic review**Preliminary review**

- Evaluate the overall effectiveness of acupuncture for osteoarthritis of the knee
- Evaluate the evidence that physical variables of acupuncture play a role in clinical outcomes

Main review

- Evaluate the evidence for warm needle vs needle acupuncture for osteoarthritis of the knee
- Gather information on clinical research protocols for the translational analysis of acupuncture for osteoarthritis of the knee (chapter-VII)

Chapter-V: Expert Interviews

- Further develop the understanding of the components of acupuncture
- Support the development a questionnaire for the practitioner survey
- Gather information on acupuncture in practice for the translational analysis of acupuncture for osteoarthritis of the knee (chapter-VII)

Chapter-VI: Practitioner survey

- Gather data on the physical components of acupuncture and associated techniques used to treat osteoarthritis of the knee in clinical practice, in the UK.
- To evaluate the intra and inter-style homogeneity and heterogeneity of acupuncture for osteoarthritis of knee

Box-II.6: Aims and objectives of phase 4**Chapter-VII: Translational analysis of acupuncture for osteoarthritis of the knee**

Aim - To develop an overview of acupuncture for osteoarthritis of the knee

Objectives

- Evaluate the external validity of the acupuncture provided in clinical trials
- Evaluate the risk of bias of sham acupuncture procedures
- Evaluate the generalisability of the evidence base
- Compare and contrast practice with theory

Chapter-VIII: Pilot study

Aim - to develop an RCT protocol to test the hypothesis:

Warm needle acupuncture leads to greater clinical benefits for those with osteoarthritis of the knee than the use of needle acupuncture alone

Objectives

- Test the integrity of the study protocol. The study will enable the evaluation of the practicalities of the procedures and identify any problems that may arise from: implementing the inclusion/exclusion criteria; patient information and consent procedures; staff training; administration of outcome assessments; randomisation and allocation concealment.
- Assess the safety of warm needle acupuncture for OA of the knee
- Assess the acceptability of warm needle acupuncture among UK patients
- Collect qualitative data from participants and staff to support the development of the protocol for an adequately powered RCT
- Provide an initial indication of the effectiveness of warm needle acupuncture compared to needle acupuncture to inform a sample size calculation for an adequately powered RCT

III. Narrative review: theoretical approaches

膝
盖
疼
痛

III.1. Introduction

This chapter is a narrative review of the theoretical approaches to the treatment of osteoarthritis of the knee with acupuncture. The original aim of this project was to develop an RCT protocol to investigate the efficacy of the additional use of moxibustion ([section-I.5](#)). The condition (bing-illness) chosen for the RCT, ideally, should have certain characteristics; so that a semi-flexible acupuncture protocol retains a high degree of external validity ([section-II.4.5.b](#)). Consequently, the key objective of this review is to assess whether osteoarthritis of the knee is a suitable condition. This judgement is made using the requirements set out in [box-II.2](#). The second objective of this review is to gather theoretical information for the translational analysis of the treatment of osteoarthritis of the knee with acupuncture ([section-II.5.4](#)). The review was used to prepare for the expert interviews ([chapter-V](#)) and develop the practitioner survey questionnaire ([chapter-VI](#)). The objectives of this narrative review are summarised in [box-III.1](#)

The review focusses on the TCM style. Whilst the researcher has knowledge of other styles his main area of expertise is TCM. The pilot study is based on TCM theory ([chapter-VIII](#)). TCM is taught in most universities and colleges in the United Kingdom, and is the dominant theory used in Chinese universities and hospitals, today. TCM has been cited as the basis for the development of acupuncture protocols in previous clinical trials (Molsberger et al., 2006). Nevertheless as there is a considerable degree of synthesis and cross-fertilisation between styles ([section-I.3](#)), the scope of the review includes other styles of practice.

Box-III.1: Objectives of the narrative review

- ▶ Evaluate whether osteoarthritis of the knee meets the criteria for the pilot study
- ▶ Gather information on theory for the translational analysis of acupuncture for osteoarthritis of the knee (chapter-VII)
- ▶ Preparation for the expert interviews (chapter-V)
- ▶ Support the development of the practitioner survey (chapter-VI)

III.2. Osteoarthritis of the knee

Osteoarthritis is defined as a group of heterogeneous conditions that lead to issues with joints, and various signs and symptoms associated with changes to the cartilage, bone and joint margins (Kumar and Clark, 2017). Diagnosis can be made using clinical examination, with or without radiographs and laboratory tests (Altman et al., 1986). Clinical diagnosis is based on the presence of pain plus at least three of the following six criteria: over 50 years of age, stiffness for less than 30 minutes, crepitus, bony tenderness, bony enlargement, no palpable warmth (American College of Rheumatology, 1986).

Ethically, it is important that the disease burden and current treatment options warrant the investigation of a novel therapy. Approximately 8.5 million UK adults have osteoarthritis; and estimates of the prevalence of symptomatic knee osteoarthritis range from 11-19% (Neogi, 2013). Globally, the disease burden of osteoarthritis in terms of disability-adjusted life years (DALYs) rose by 64% from 2000-2010 (Murray et al., 2012). With an ageing population and an increase in obesity, it is likely that the prevalence of osteoarthritis of the knee will continue to rise (Conaghan, 2008). The public health impact of osteoarthritis is substantial, at both the individual and societal level (Neogi, 2013). Osteoarthritis is among the top ten causes of disability worldwide (Vos et al., 2012). Current guidelines for the treatment of osteoarthritis include exercise, non-steroidal anti-inflammatory drugs (NSAIDs) and surgery (NICE, 2014b).

However, the therapeutic effects of pharmacological interventions for osteoarthritis are small and limited to the first few weeks after treatment (Bjordal et al., 2007, Bjordal et al., 2004). Furthermore, NSAIDs can cause serious gastro-intestinal side effects, including dyspepsia, peptic ulceration, and haemorrhage, and even result in death (Hippisley-Cox et al., 2005). The annual direct and indirect costs per individual were estimated to be 5,700 US dollars in 2000 (Neogi, 2013). In the UK the estimated costs in 2010 were £850 million for hip and knee replacement, £25 million for oral non-steroidal anti-inflammatory drugs (NSAIDs), and £56.9–£124.4 million for iatrogenic events related to NSAID use (Chen et al., 2012). This suggests that there is a need for better treatment options.

III.3. Methods

Due to the diversity of TBSA ([section-I.3](#)) it was necessary to establish some limitations. It was also decided that the review would concentrate on modern textbooks rather than classical texts. The premise of this decision was the belief that modern textbooks better reflected the theoretical understanding of the majority of practitioners in the UK. The intention was to select books that had been widely used and were familiar to many practitioners. In addition, a balance was sought on a number factors. Books that have been in use for some time as well as recent publications were reviewed. Practitioners who qualified over 20 years ago might not use recently published textbooks, recent graduates may not have the older publications. Practitioners' understanding of the theory is probably strongly influenced by the texts that were available when they were students. Books that were written by Chinese authors, published in China, were balanced with those by non-Chinese, published in the West. The UK has ethnic Chinese practitioners who initially trained in China. Some Chinese language textbooks were also reviewed to account for the learning experiences of these

practitioners. The researcher is able to read modern Chinese, so no translation was required.

The reviewed books are set out in [tables III.1&2](#). Once these texts had been examined, it was felt that saturation had been reached and little would be gained by extending the review. A total of nine TCM textbooks were reviewed ([table-III.1](#)), three Japanese, five Five-element, and three medical acupuncture ([table-III.2](#)).

As a course director of an acupuncture programme, with external examining experience, the researcher is familiar with the field. To guard against bias the principals of three other colleges (one TCM, one Five-element, one Stems and branches) were contacted for advice, as was a teacher of Japanese acupuncture and two Medical acupuncture educationalists. Discussions were also conducted with colleagues at London South Bank University, who had specific knowledge of Japanese or Five-element acupuncture, or had trained in China.

Table-III.1: Reviewed textbooks - TCM Style

Author	Title	Language
Legge and Vance, 1997	Close to the bone	English
Maciocia, 1994	The practice of Chinese medicine	English
O'Connor and Bensky, 1981	Acupuncture : a comprehensive text	English
Cheng, 1999	Chinese acupuncture and moxibustion (Revised Edition)	English
Xu et al., 1988	Acupuncture treatment of common diseases	English
Zhao, 2002	Chinese acupuncture and moxibustion	English/ Chinese
Zhao and Wang, 2007	Chinese medicine study guide: Acupuncture and moxibustion	English
Sun, 2000	The study of acupuncture	Chinese
Wu and Sheng, 2002	The study of practical acupuncture	Chinese

Table-III.2: Reviewed textbooks - non-TCM styles

Author	Title	Style
Matsumoto and Euler, 2008	Kiiko Matsumoto's Clinical Strategies	Japanese
Shudo, 2003	Finding effective acupuncture points	Japanese
Birch and Ida, 1998	Japanese acupuncture: A clinical guide	Japanese
Hicks et al., 2010	Five Element Constitutional Acupuncture	Five element
Jarrett, 2006	The clinical practice of Chinese medicine	Five element
Jarrett, 1999	Nourishing Destiny	Five element
Worsley, 1998	The Five elements and the officials	Five element
Worsley, 2012	Traditional Diagnosis	Five element
Golding, 2013	The Complete Stems and Branches	Stems and Branches
White et al., 2008	An introduction to Western medical acupuncture	Medical
Ma et al., 2005	Biomedical acupuncture for pain management	Medical
Baldry, 2004	Acupuncture, Trigger Points and Musculoskeletal Pain	Medical

III.3.1. Selection of TCM textbooks

Maciocia (1994) is an extremely well known author whose books have been used by most, if not all, TCM colleges for some time. 'Close to the bone' focuses on musculoskeletal problems, it is predominantly a TCM text but includes a biomedical perspective (Legge and Vance, 1997). After the Reform and Opening up Policy (改革开放)⁶ circa 1978, Western people were able to go to China to study acupuncture. Three international acupuncture training centres;

⁶ These reforms brought an end to China's decades of isolation

Beijing, Shanghai and Nanjing were established. 'Chinese acupuncture and moxibustion' was first published in 1987 and was used as a core text used by these training centres (Cheng, 1999). 'Acupuncture treatment of common diseases' was produced by the international training centre in Nanjing (Xu et al., 1988). These books have also been used by colleges in the UK. More recent publications from these international training centres were also reviewed (Zhao and Wang, 2007, Zhao, 2002) 'Acupuncture: a comprehensive text' was one of the first books published by Western authors after the opening-up policy (O'Connor and Bensky, 1981). 'The study of acupuncture' (Sun, 2000) is a widely used text in China. 'The study of practical acupuncture' (Wu and Sheng, 2002) is a Chinese language text compiled by practitioners from Nanjing, many of whom taught Westerners who then went on to become educationalists in the UK.

III.3.2. Selection of non-TCM textbooks

TCM is an amalgamation of different schools of practice (Birch and Felt, 1999, Scheid, 2002). As it has been through a process of systemisation there is a degree of homogeneity. There is greater heterogeneity when considering non-TCM styles. Because of the diversity of TBSA ([section-I.3](#)) the size of the review could have become unmanageable. Surveys have indicated that Five-elements is the second most common TBSA after TCM; Japanese and other styles were practiced by less than 10% of practitioners (British Acupuncture Council, 2011, Hopton et al., 2012). Colleges that have accreditation by the British acupuncture accreditation board (BAAB) typically state that they teach one or more of the following styles: TCM, Five-elements, or Stems and Branches (BAAB, No date). Therefore, it was decided to only review textbooks of the Five-elements, Japanese, and Stems and Branches styles as well as medical acupuncture. This meant that apparently less common styles such as Korean acupuncture or Master Tung were not reviewed.

Both Japanese-style practitioners who were consulted recommended all three textbooks (Birch and Ida, 1998, Matsumoto and Euler, 2008, Shodo, 2003). Worsley was the founder of the five-element style in the UK, therefore his books

were reviewed (Worsley, 1998, Worsley, 2012). The authors of 'Five Element Constitutional Acupuncture' (Hicks et al., 2010) are principals of one of the main colleges. Jarrett's books (Jarrett, 2006, Jarrett, 1999), were selected on recommendation.

III.3.3. Methods of data extraction and analysis

The review of TCM textbooks was conducted first. Tables were compiled on the aetiology and differential patterns, sign and symptoms, treatment methods and recommended acupuncture points (appendices-B.1-6). Textbooks of other styles of practice were then examined and analysed in relation to the information regarding TCM. Confirmation of key findings regarding non-TCM styles of practice was sought from the practitioners/ educationalists who had helped identify the textbooks.

III.4. Basic acupuncture theory

There are two components in the TBSA diagnostic process - illness differentiation *bianbing* 辨病 and pattern differentiation *bianzheng* 辨证 (section-I.3.5) For those readers who are not familiar with TBSA theory, further explanation of these key concepts is given in this section. In addition, the 'mechanism' of pain within Chinese medicine theory is also described.

Bianbing

The nosology of TBSA and biomedicine are different, as such it is not possible to identify a *bing-illness* 病 that exactly corresponds with osteoarthritis of the knee. The *bing-illness* that bears the most similarities is *xi tong* 膝痛: literally 'knee pain'. However, this category includes knee pain from other causes, such as injury. Osteoarthritis of the knee may potentially be classified as Crane's knee, a situation where there is deformation of the joint and muscle wasting above and below the knee.

Bianzheng

Pattern differentiation guides the selection of acupuncture points and the other physical processes of treatment ([section-I.3.5.b](#)). Thus, knowledge of the potential differential patterns associated with ‘knee pain’ is required to understand how osteoarthritis of the knee would be treated. The signs and symptoms typically associated with osteoarthritis of the knee would be classed as *bi-zheng* 痹症 bi-syndrome.

III.4.1. Bi-syndrome

Bi-syndrome, like osteoarthritis, can affect different joints within the body. Bi-syndrome, however, encompasses a wider range of conditions. It can effect the muscles as well as joints. Biomedical conditions such as gout, fibrositis, rheumatoid arthritis and sciatica can also be considered bi-syndrome (Cheng, 1999, Wu and Sheng, 2002, Xu et al., 1988). It is probably the case that all patients with osteoarthritis would be differentiated as having bi-syndrome, but not all cases of bi-syndrome would be diagnosed as osteoarthritis.

III.4.2. Pain ‘mechanism’ in TBSA

The word *bi* 痹 is a homonym of *bi* 闭; which means blockage or obstruction. Chinese language textbooks emphasise that the meaning of *bi* 痹 is *bi* 闭, namely blockage/ obstruction (Sun, 2000, Wu and Sheng, 2002). The syndrome has been translated as painful obstruction syndrome (Maciocia, 1994) or blockage syndrome (Chen and Shun, 2005). ‘Blockage’ causes the pain, in fact all pain is considered to be the result of blockage. This is illustrated by the expression:

通则不痛，痛则不通

tong ze bu tong, tong ze bu tong

Where there is free flow [of Qi] there is no pain

Where there is pain there is no free flow [of Qi]

Consequently, in TBSA the ‘mechanism’ of pain is relatively straightforward and simple. This simplicity was considered a beneficial trait for RCT protocol ([chapter-VIII](#)). See [section-III.5.5.a](#) for further explanation.

III.4.3. Root and Branch - Ben and Biao

The pattern differentiation process typically involves identifying two key components, the *ben* (本) and the *biao* (标). *Ben* is often translated as 'root' and *biao* as 'branch'. The difference between the two is nuanced, but for simplicity consider the branch as the relatively external parts of the body and the root as the internal.

In knee pain, the actual pain, swelling and stiffness are considered to derive from a blockage within the meridians. The meridians are by definition external. Therefore, the signs of symptoms that relate specifically to the knee are used to assess **the branch**; for example, the location of the pain, the nature of the pain, the factors that ameliorate the pain, any discolouration of the skin, stiffness or swelling.

The root on the other hand refers to the internal functioning of the body. The pattern differentiation of the root is made by assessing the general health of the patient. For example the digestion, sleep patterns, speed of the pulse and shape of the tongue are all relevant signs and symptoms. TBSA posits that factors such as age, quality of sleep, and digestion, can all influence the body's ability to nourish the knee joints. The pattern differentiation of osteoarthritis of the knee may include: an analysis of the 'blockage in the meridian' that is causing the pain (the branch); an assessment of the internal factors that may allow the blockage to develop and/or persist (the root).

III.5. Results - The standard theory

There is a common theory shared by all the TCM textbooks reviewed. This is referred to as the 'standard theory' in this thesis. The 'standard theory' is described in this section. The protocol of the pilot study is based on the standard theory ([chapter-VIII](#)). Some of the TCM textbooks discussed an

alternative theory. The alternative TCM theory and the theories of other styles of practice are presented in [section-III.6](#).

III.5.1. Standard theory - The branch

The standard theory conceptualises the aetiology of bi-syndrome as an external invasion of pathogenic factors: Wind, Cold and Damp ([table-III.3](#)). These pathogens, Wind, Cold and/or Damp, are said to attack the meridians *jing* 经 and collaterals *luo* 络.⁷ The meridians and collaterals are by definition superficial. Therefore, blockages in the meridians and collaterals are considered part of the branch in the differential pattern.

There are three primary sub-patterns, that are differentiated from one another by the characteristic of the pain. Each sub-pattern has two names. Either named after a key characteristics of the pain or alternatively the invading pathogen that is thought to predominate ([box-III.2](#)). The sub-categories can be combined. For example, if a patient complains of very severe pain that is made worse by damp weather this could be identified as Cold Damp Bi syndrome.

The textbooks are in general agreement regarding the signs and symptoms that identify each sub-pattern ([table-III.4](#)). Some signs and symptoms used to identify the sub-patterns are not related directly to the knee. Typically, these are the tongue and pulse indications ([table-III.5](#)). The full version of [tables-III.3-6](#) can be found in ([appendices-B1-4](#)).

⁷ Acupuncture charts typically depict only the meridians. However according to the theory there is also a vast network of smaller vessels known as *luo* which covers the entire surface of the body.

Box-III.2: Sub-patterns of bi-syndrome - Standard theory

Moving Bi (*xing bi* 行痹): soreness and pain in the joints, limitation of movement with pain moving from joint to joint. The key characteristic is that the pain moves from joint into joint. Moving Bi occurs when Wind predominates and can also be called **Wind Bi** *feng bi* 风痹.

Fixed Bi (*zhuo bi* 着痹): pain, soreness, swelling of the joint, feeling of heaviness or numbness, worse in damp weather. The alternative name of syndrome is **Damp Bi** *shi bi* 湿痹 and occurs when Damp predominates. The key characteristic of this subcategory is that the pain is fixed and does not move from joint to joint.

Painful Bi (*tong bi* 痛痹): severe pain in the joint or muscle with limitation of movement, worse in cold weather. The alternative name is **Cold Bi** *han bi* 寒痹 syndrome. The key characteristic is that this subcategory is that it is especially painful.

Heat Bi (*re bi* 热痹): pain, feeling of heat in the joint, joint feels hot to touch, swelling and redness.

Within the standard theory there is another form of Bi syndrome - Heat Bi. This subcategory is commonly seen as a progression from one of the first three subcategories. In TBSA there is an expectation that the pattern differentiation will change over time ([section-I.3.5.b](#)). If the illness is not resolved, the obstruction leads to the generation of heat and the symptom picture changes, resulting in a modification of the differential pattern.

Maciocia (1994) adds a fifth subcategory, bone painful obstruction syndrome. The key sign is that the bones are visibly deformed. This probably refers to the Chinese term *wang bi* 尙痹 deformity Bi. However, as the term *gu bi* 骨痹 Bone Bi is used in another theory of bi-syndrome classification ([section-III.6.1](#)) there is an element of ambiguity, not untypical of TBSA. O'Connor and Bensky (1981) identify another possible sequelae, Heart Bi, with symptoms such as discomfort in the chest and palpitations.

III.5.2. Standard theory - The root

With the exception of O'Connor and Bensky (1981), all TCM sources state that the pathogenic factors are allowed to enter the body due to an internal deficiency. To be more precise, a relative deficiency. This relative deficiency can be considered the root differential pattern. Reasonably healthy individuals can be affected, if they are exposed to severe Wind, Cold and/or Damp. On the other hand an individual could be affected following a mild exposure, if their general health is weak. The relative weakness/strength of the external pathogen and general health is evaluated by assessing all the signs and symptoms.

The descriptions of the exact nature of this weakness, its cause, and relative importance, give rise to an apparent heterogeneity between the textbooks. In general, the root is seen as a deficiency of Yang defensive Qi, that allows the pathogenic factor to enter the space between the skin and muscle known as *couli* (Cheng, 1999, Maciocia, 1994, Sun, 2000, Wu and Sheng, 2002). Some textbooks describe the weakness as being due to deficiency of Blood or Yin, that lead to a malnourishment of the meridians (Legge and Vance, 1997, Maciocia, 1994). Alternatives include a deficiency of Liver and Kidney, obstruction due to Phlegm with Spleen deficiency or long-term Stagnation of Blood in the joints (Zhao, 2002). Most authors focus on the external invasion, the branch. Zhao (2002) and Wu and Sheng (2002) emphasise the *internal* root. Legge note the relative importance of treating the root may vary from patient to patient (Legge and Vance, 1997).

Table-III.3: Standard theory - mechanisms and aetiology (abridged)

Author	Mechanism	Further notes	Aetiology
O'Connor and Bensky, 1981	<ul style="list-style-type: none"> Invasion of Wind, Cold, Damp 	<ul style="list-style-type: none"> Heat Bi - Wind Cold Damp Bi leads Heat Can effect organs typically leads to Heart Bi 	
Cheng, 1999	<ul style="list-style-type: none"> Invasion of Wind, Cold, Damp General weakness of the body and deficiency of Yang Qi Dysfunction of the pores and weakness of defensive Yang 	<ul style="list-style-type: none"> Heat Bi - If body constitution is Yang can lead to Heat Heat Bi - Long-standing Wind Cold Damp Bi leads to Heat Alternative method of classification, Skin Muscle, Tendon, Vessel, Bone Bi 	<ul style="list-style-type: none"> Common in areas where the weather is cold, wet and windy Affects both sexes Any age
Xu et al., 1988	<ul style="list-style-type: none"> Invasion of Wind, Cold, Damp Body is weak Body surface is loose 	<ul style="list-style-type: none"> Heat Bi - Wind Damp Bi can turn to Heat Prolonged Bi can move inward to affect the organs 	
Zhao, 2002	<ul style="list-style-type: none"> Obstruction by Wind, Cold, Damp, Heat Deficiency of Liver and Kidney Obstruction by Phlegm due to Spleen Deficiency Long term stagnation of Blood in the joints 		<ul style="list-style-type: none"> Elderly more likely to be affected
Zhao and Wang, 2007	<ul style="list-style-type: none"> Invasion of Wind, Cold, Damp Deficient Root Excess Branch 	<ul style="list-style-type: none"> Prolonged stagnation transforms into Heat (apparent typing error in the flow chart Page 239) 	

Table-III.4: Standard theory - signs and symptoms related to the knee (abridged)

Author	Wind Bi	Damp Bi	Cold Bi	Heat Bi
Legge and Vance, 1997	Soreness in muscles and joints, Pain will move from place to place, pain is varying in intensity, Decreased range of movement	Soreness and heavy sensation of a joint or muscles, numbness in skin over the affected area, pain does not move pain worse in cloudy and wet weather	Pain is the most prominent symptom, no inflammation, pain is fixed in the joint, better for warmth, worse for cold, decreased range of passive movement	Joint pain with redness and swelling, Limitation of movement,
Maciocia, 1994	Soreness and pain of the muscles and joints; limitation of movement; with the pain moving from joint to joint	Pain, soreness and swelling in muscles and joints with a feeling of heaviness and numbness of the limbs, the pain being fixed in one place and aggravated by damp weather	Severe pain in a joint or muscle with limitation of movement, usually unilateral	Pain and heat in the joints which feel hot to the touch, redness and swelling of the joints, Limitation of movement and severe pain
O'Connor and Bensky, 1981	Pain in joints is widespread, moves from one area to another	The pain is localised and does not move. The body and limbs are heavy, there is numbness, oedema	Severe pain one part, all over one half of the body. Worse when exposed to cold, better for warmth.	The flesh is hot, the area of pain is red and swollen. Pain increases on contact
Cheng, 1999	Pain in the joints, especially the wrists, elbows, knees and ankles; limitation of movement	Heavy sensation of the limbs, soreness and fixed pain in the joints, aggravated on cloudy and rainy days	Severe stabbing pain in the joints, Alleviated by warm and aggravated by cold, with fixed localisation but to no local redness and hotness	Arthralgia involving one or several joints, local redness, Swelling and excruciating pain with limitation of movement
Xu et al., 1988	Pain moving from one area of the body to another, or radiating to a remote area accompanied by pulling sensation and numbness,	Soreness, Pain and heaviness of the limbs; may be induced or aggravated by cloudy, windy, Cold or rainy weather; Slight swelling of the muscles in skin, absence of local redness	Intense pain and a cold sensation; the pain is localised and may worsen with cold and diminish with warmth	Pain, redness, swelling and a burning sensation of the joints, Which are tender on contact and have restricted movement

Table-III.5: Standard theory - additional signs and symptoms (abridged)

Author	Wind Bi	Damp Bi	Cold Bi	Heat Bi
Legge and Vance, 1997	Sometimes accompanied by local heat and chills, superficial and rapid pulse, been white sticky timecode	Thick white and sticky tongue coat, soft slow pulse	Thin white tongue coat, deep wiry tense pulse	Fever, thirst, yellow tongue coat and rapid rolling pulse
Maciocia, 1994	Acute cases the pulse floating and slightly rapid	Acute cases the pulse would be slow and slightly slippery	Acute cases the pulse is tight	Acute cases there would be thirst fever which does not abate after sweating and a slippery rapid pulse
O'Connor and Bensky, 1981	There is often fever and chills, rapid pulse and yellow fur on the tongue	Greasy fur on the tongue, moderate pulse	The tongue fur is thin and white, pulse wiry and tight	Mouth and tongue are parched, urine dark, constipation. tongue fur is yellow and greasy, rapid pulse
Cheng, 1999	Chills and fever, Fingers sticky tongue coating, Superficial and tight or superficial and slow pulse	White and sticky tongue coating and soft pulse	Thin and white tongue coating, string taut and intense pulse	Fever and thirst, yellow tongue coating, Rolling and rapid pulse
Xu et al., 1988	The time coating is yellow or light yellow and thin; pulse is superficial and wiry; in some cases chills and fever may also be present	Soreness, Pain and heaviness of the entire body; Are white and sticky tongue coating and a softer pulse	There may be a white tongue coating and attends wiry polls	Symptoms may include fever, sore throat and profusely sweating which does not reduce body temperature. There is also a yellow and sticky tongue coating and a rapid pulse

Table-III.6: Standard theory - treatment methods

Author	General	Wind Bi	Damp Bi	Cold Bi	Heat Bi
Legge and Vance, 1997	Herbs	<ul style="list-style-type: none"> relatively more points are used along the channel reducing method don't retain the needles treat every day 	<ul style="list-style-type: none"> Used mainly local and A-shi points Distal points not necessary for every treatment Direct moxibustion and warm needle Or needles only 	<ul style="list-style-type: none"> Acupuncture and moxibustion Retain needles for 20 minutes Treat every second day 	<ul style="list-style-type: none"> Prick congested veins then cupping Reducing method Treat every day
O'Connor and Bensky, 1981	<ul style="list-style-type: none"> local and distant points on channels which traversed the area of pain Severe cases treat once-a-day, for most cases treat on alternating days Cupping with cutaneous acupuncture 	Needling is the primary method	Combination of needling and moxibustion	Combination of needling and moxibustion	Let a few drops of blood at related points
Xu et al., 1988	<ul style="list-style-type: none"> Removal obstruction Regulate circulation of Qi and Blood Select local and distal points on meridians that traverse the area A-shi points can also be selected <p>Treatment every other day or daily in severe cases</p>	<ul style="list-style-type: none"> Disperse Wind Reducing method Shallow needling 	<ul style="list-style-type: none"> Disperse Cold Needles retained for longer Moxibustion - emphasised 	<ul style="list-style-type: none"> Resolve Damp Needles and moxibustion 	<ul style="list-style-type: none"> Clear Heat Reducing method Shallow needling Pricking method

III.5.3. Replicability of the pattern differentiation

One of the objectives of this review is to assess whether osteoarthritis of the knee is a suitable condition for the pilot study; i.e. a semi-flexible acupuncture protocol would retain a high degree of external validity ([section-II.5.2](#)). The criteria are: a small number of associated differential patterns; minimal differences between the points recommended for these patterns; key diagnostic signs and symptoms that are likely to have high inter-rater reliability ([section-II.4.5](#)). In addition, it is likely that all participants entering a study on the basis that they had osteoarthritis of the knee would be considered to have bi-syndrome

III.5.3.a. Replicability of the branch

A clinical diagnosis of osteoarthritis of the knee requires the presence of at least three of six criteria, in addition to pain ([table-III.7](#)) (American College of Rheumatology, 1986). Apart from 'age above fifty years', all the signs and symptoms are related directly to the knee. As a result, they are associated with the branch not the root.

In osteoarthritis the pain does not move from joint to joint, therefore, it is unlikely to be considered Moving Bi. Similarly, it is unlikely to be classified as Heat Bi syndrome as there should be no palpable warmth ([table-III.7](#)). The signs of bony tenderness or enlargement are not part of the initial stages of bi-syndrome but may appear as the condition progresses (Legge and Vance, 1997, Maciocia, 1994). Crepitus or above 50 years old, whilst not listed in the textbooks, would not preclude an individual from being classified as having bi-syndrome.

Summary, osteoarthritis of the knee would typically be classified as Painful or Fixed Bi, also known as Cold or Damp Bi. As the condition progresses it might transform into Bony Bi. The small number of potential sub-patterns suggests that a replicable protocol would have a high degree of external validity ([box-II.2](#)).

Table-III.7: Comparison of knee osteoarthritis and bi-syndrome

ACR	Painful Bi	Fixed Bi	Moving Bi	Heat Bi
pain	✓	✓	X	✓
stiffness	✓	✓	✓	✓
bony tenderness	✓	✓	✓	✓
bony enlargement	—	—	—	—
crepitus	—	—	—	—
> 50 years	—	—	—	—
No palpable warmth	✓	✓	—	X

ACR -American College of Rheumatology (diagnostic criteria)

Pattern differentiation is a subjective process ([section-I.3.5.b](#)). This subjectivity starts with the assessment of signs and symptoms. Certain signs and symptoms have a greater inter-rater reliability than others ([section-II.3.3](#)). Assessment of knee pain has been shown to have ‘almost perfect’ inter-rater reliability; stiffness and swelling were ‘fair’ (Hua et al., 2012). This suggests a high degree of external validity could be achieved by a replicable protocol.

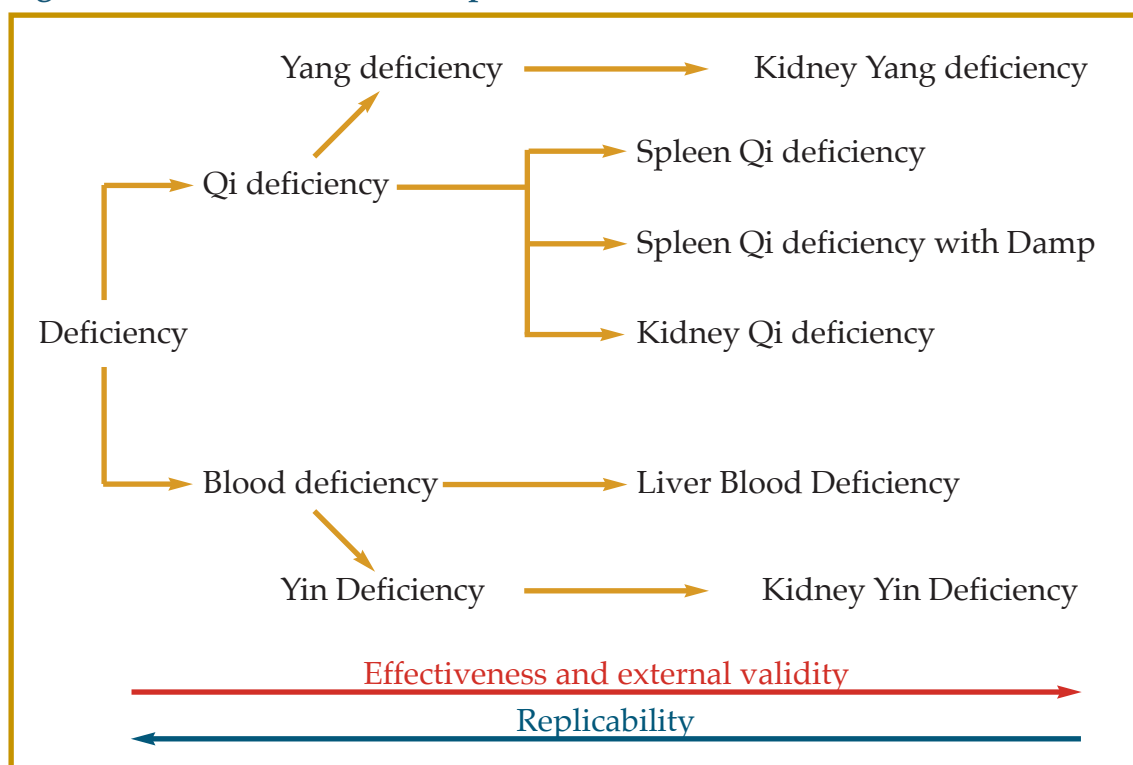
III.5.3.b. Replicability of the root

Pattern differentiation of the root is based on a wider variety of signs and symptoms, such as the tongue and pulse, some of which may have poor inter-rater reliability ([section-II.3.3](#)). Within the standard theory there greater heterogeneity in the possible scenarios for the root compared to the branch. However, these are not necessarily contradictory ideas that reflect fundamental differences in opinion amongst authors. Indeed, Maciocia presents both Qi deficiency and Blood deficiency (1994).

Pattern differentiation can be conceptualised as a process of refinement ([section-II.4.5](#)). Qi or Blood deficiency are broad brushstrokes which can be further refined. The reasons why an individual may have Qi and/or Blood deficiency are numerous (Maciocia, 2005). Typically, a complete pattern differentiation of the root would also identify problems with individual organs. Deficiency of

Liver and Kidney are elaborations on Blood/Yin deficiency; Spleen deficiency is an elaboration on Qi deficiency (Maciocia, 2005, Zhao, 2002). An exhaustive list scenarios would require a description of all the possible combinations of signs and symptoms that may accompany osteoarthritis of the knee, such as digestion, sleep, the colour of the tongue, the pulse characteristics, headaches and so on. This is not be feasible. Authors can only highlight the scenarios that are the most common in practice. There is an assumption that the practitioner knows how to refine the differential pattern. In this respect Qi and/or Blood deficiency should be seen as common scenarios, rather than definitive components of the pattern differentiation. In short, the apparent differences may be due to editorial decisions, rather than actual disagreements on the theory.

Figure-III.1 illustrates the refinement of the root pattern differentiation using the standard theory. The first step is to decide whether or not the patient is deficient. If so, the next step is to assess whether Qi or Blood deficiency predominated. If Qi deficiency, then this may be related to the Spleen or Kidney or involve Damp. At each stage there is the potential for practitioners to interpret the signs and symptoms differently. These decisions are dependent on the training and ability of the practitioners. External validity and theoretically effectiveness increases as one goes through each stage - moves from left to right in figure-III.1. Conversely, replicability decreases.

Figure-III.1: Illustration of root pattern differentiation

III.5.3.c. Replicability - overall pattern differentiation

The pattern differentiation of the branch appears to be more amenable to replication than the root. The inclusion the root differentiation marks a shift from a stage 2 to a stage 3 protocol (section-II.4.5). Theoretically, this would lead to a more effective and externally valid treatment, but at the expense of replicability (figure-II.9). It is possible that practitioners do not consider treating the root particularly important for certain conditions (section-II.4.5). Most of the textbooks focus on the branch pattern differentiation, only one concentrated on systemic points and did not suggest local or meridian points⁸ (Wu and Sheng, 2002). If treatment of the root is not considered important in practice, then a semi-flexible protocol will retain a high degree of external validity; and meet the criteria for the pilot study (chapter-VIII). However, it is not possible to

⁸ This book was compiled by practitioners affiliated to Nanjing University of Chinese medicine. The researcher spent a year in Nanjing working with some of the authors who in practice use local points around the knee.

gauge how important the root is in practice from this review. This assessment can only be made using data from the practitioner survey ([chapter-VI](#)).

III.5.4. The standard theory - Ageing

The sole criterion of clinical osteoarthritis that does not pertain directly to the knee was 'age greater than 50'. In TBSA the ageing process is linked to Shen Xu (肾虚), commonly translated as Kidney deficiency. It is important to note this does not equate to a kidney condition in biomedicine. Weakness or problems with the knees are signs of Kidney deficiency (Maciocia, 2005). The Kidney Qi is thought to gradually decline as a person gets old. The implication is that an internal deficiency is, to some extent, inevitable. This has consequences for the prognosis. As the condition is linked to the ageing process then treatment may be directed at ameliorating the signs and symptoms rather than a complete 'cure'.

The effects of ageing can be mitigated by a good diet, exercise and rest (Maciocia, 2005). Acupuncture and herbal medicine are also thought to be able to help maintain general health, which in turn means that the body is better able to nourish the knee joints. TBSA treatment for a person with Shen Xu and knee pain may involve lifestyle advice as well as acupuncture treatment. Adherence to the lifestyle advice may impact on the long-term outcome.

III.5.5. Replicability of treatment methods

This project seeks to understand acupuncture in terms of its physical components ([section-II.2.2](#)). [Figure-II.5](#) illustrates three subjective stages: assessment of signs and symptoms; pattern differentiation; selection of acupuncture points and auxiliary methods. The first two are discussed above ([section-III.5.3](#)). The final stage leads to the objective physical process of acupuncture.

There are two issues of concern. This review indicates that osteoarthritis of the knee would normally be associated with only three differential patterns: Cold Bi, Damp Bi and Bony Bi ([section-III.5.3.a](#)). If the suggested treatment methods

for each of these sub-patterns are radically different, then the reproducibility of the protocol would be reliant on good inter-rater correlation between practitioners for pattern differentiation ([section-II.3.3](#)). The second issue is the degree of heterogeneity among authors. In other words, do different authors suggest different treatment methods for the same sub-pattern. Both of these issues could limit the generalisability of the pilot study ([chapter-VIII](#)).

III.5.5.a. Modes of acupuncture point selection

The recommended acupuncture points are presented in ([appendices-B.5/6](#)). However, to make the data more meaningful it is necessary to group the acupuncture points according to the rationale of the standard theory. In other words, apply component analysis to the clinical reasoning process of point selection. The point selection process can be reduced to four components. In this thesis these components are described as '*modes of acupuncture point selection*'.

The four modes of point selection are:

- i. Local points
- ii. Meridian: selected according to the meridian that traverses the area of pain
- iii. General health: commonly recommended points, for this condition, to support the general health of the patient
- iv. Differential pattern points: points selected as part of an individualised treatment based on the ben-root pattern differentiation

Local and meridian points

In TBSA pain develops from blocked Qi ([section-III.4.2](#)). In the standard theory the first objective of treatment is to unblock the affected meridian. There is wide agreement that the key to treatment is the selection of local points and points along the affected meridian ([appendix-B.4](#)). Thus, the key aspect of pattern differentiation is to locate the pain. If one assumes that patients would consistently indicate the painful locations, then similar local and meridian points ought be selected by different practitioners. This implies a high degree of

inter-rater reliability in the selection of the most important acupuncture points.

There are two caveats. Firstly, the research on inter-rater reliability of signs and symptoms of osteoarthritis of the knee did not include the detail of assessing which meridian was affected (Hua et al., 2012). Secondly, there may be more skill involved in locating the most tender point than anticipated. Ideally, these questions could be addressed in future studies.

There are some differences in the choice of local and meridian points depending on the differential pattern. The use of a-shi points, places painful on pressure, is emphasised in the treatment of Cold Bi syndrome (Legge and Vance, 1997). Whereas points along the course of the meridian are highlighted for Wind Bi (Legge and Vance, 1997). However, for Cold Bi and Damp Bi the points are similar. Consequently, even if practitioners disagree on the predominant pathogen, Cold or Damp, the local and meridian points should still be similar.

General health and underlying root pattern points

Acupuncture points that are thought to address the underlying root pattern are also recommended in the textbooks. These varied depending on the subcategory of bi-syndrome ([appendix-B4](#)). There is a reasonable amount of agreement among authors regarding some of these points. For example, shenshu BL23 and guanyuan Ren4 were identified by seven out of nine books. These can be considered generic points to help support the general health of a 'typical' patient with osteoarthritis. Some points are only recommended by one textbook ([appendix-B.6](#)). The heterogeneity is in part related to the different scenarios of the root differential patterns ([section-III.5.2](#)). Just as it is not possible to set out all possible differential patterns, it is also not feasible to produce an exhaustive list of all possible points. Some authors refer to additional points suggested in classical texts (Sun, 2000). There is considerably greater heterogeneity of the underlying root pattern (distal) points compared to the local and meridian points ([figure-VII.2-3](#)).

III.5.5.b. Other physical variables

Beyond the selection of acupuncture points there are other physical variables, which may also lead to heterogeneity. These relate to the ways in which the points are stimulated, auxiliary techniques, and the course of treatment.

Needle retention time and depth of insertion

Some sources recommend that these variables should be adjusted depending on the sub-category. For Cold or Damp Bi the needles should be inserted more deeply and retained for longer, whereas for Wind Bi shallow insertion without retention is suggested (Legge and Vance, 1997, Zhao and Wang, 2007). Shallow insertion is also advocated when the illness location *bingwei* 病位 is in the skin, and deep insertion when the problem is in the bones (Cheng, 1999, Sun, 2000). Pricking and plum blossom needle are recommended for Heat and Wind Bi ([appendix-B.4](#)) but can also be used for Damp Bi (Maciocia, 1994). Intra-dermal needles are suggested for Cold Bi (Cheng, 1999). Importantly, potential inter-rater disagreement regarding whether Cold or Damp predominate would have minimal impact on the depth of insertion and needle retention time.

Needle manipulation techniques

TBSA include the notion that by manipulation of the needle the Qi can be strengthened, reduced or simply stimulated. (Cheng, 1999, Sun, 2000). These are called tonifying, reducing and even techniques, respectively. There are a number of different physical methods to achieve this aim - e.g. some rotate needle in different directions, others use speed of insertion/withdrawal (Cheng, 1999). A number of sources indicate that tonification, even or reducing technique should be used, however, without further elaboration it is not possible to know which method of physically moving the needle is being advocated.

Moxibustion

Moxibustion is indicated for Cold Bi and Damp Bi syndrome - different methods are recommended, including warm needle acupuncture and direct moxibustion. For Cheng (1999) moxibustion should be considered the primary

mode of treatment for Cold Bi with needles taking a supporting role. A number of sources indicate that moxibustion should not be used for Heat Bi (Maciocia, 1994, Sun, 2000).

Cupping

Cupping, although not commonly cited, is recommended. It can be used in combination with plum blossom needle for Damp Bi (Maciocia, 1994, Zhao, 2002).

Course of treatment

The course of treatment is not discussed by most authors. Those that do provide some information suggest treatment should be delivered once a day or every other day. In severe cases long-term treatment may be required (Legge and Vance, 1997, Xu et al., 1988).

Advice

Potential lifestyle advice involves avoiding some of the behaviours that are thought to cause the condition: including dietary advice and injunctions to keep the affected joint warm (Legge and Vance, 1997, Maciocia, 1994, Sun, 2000, Wu and Sheng, 2002).

III.5.6. Summary - Standard theory

Within the standard theory, osteoarthritis of the knee is likely to be classified as Cold or Damp Bi syndrome or the combination of the two, Cold Damp Bi. These might progress to become Bony-Bi. As there are comparatively few associated differential patterns this indicates that osteoarthritis of the knee should meet one of the key requirements of the RCT ([section-III.5.3.a](#)). In addition, inter-rater reliability should be high for the key signs and symptoms, the location and nature of the pain, that guide the selection of the main points ([section-III.5.3.a](#)). Even if practitioners discern patterns, for a particular patient, the actual points selected may be similar: as there is considerable overlap in the recommended points for each sub-pattern ([section-III.5.5](#)). All of these factors suggest that a semi-flexible protocol will have a high degree of external validity. Therefore, osteoarthritis of the knee is a suitable condition for the RCT.

One of the aims of the project is to investigate the additional use of moxibustion. This review shows not only that moxibustion is indicated for osteoarthritis of the knee, but is considered an important part of treatment. Warm needle acupuncture is one of the recommended methods of using moxibustion ([section-III.5.5](#)).

Pattern differentiation of the underlying root can only be incorporated into a semi-flexible protocol to a limited extent. There is agreement regarding some potential points. However, authors describe different scenarios for the underlying root patterns and there is a heterogeneous range of suggested points. This heterogeneity suggests a semi-flexible and repeatable protocol would compromise external validity ([section-III.5.3.b](#)). Therefore, the relative importance of treating the root is a significant issue ([III.5.3.c](#)). If the root is considered of minimal importance, then the compromises should have a minimal impact. The practitioner survey ([Chapter-VI](#)) seeks to understand whether treatment of the root is commonly addressed in practice.

III.6. Results - Alternative theories

III.6.1. Alternative TCM Bi-syndrome theory

Some TCM authors describe a second theoretical approach. This approach is related to Five-element theory. Different styles such as TCM and Five-element have a synthetic relationship (section-I.3), therefore finding an overlap is not unexpected. Within the alternative theory the tissues of the body are linked to the internal organs. As the pathogenic factor progressively moves deeper into the body signs and symptoms related to internal organs are produced (box-III.3) (Chen and Shun, 2005, Maciocia, 1994).

Box-III.3: Sub-patterns of bi-syndrome - Alternative TCM theory

Skin Bi (*pi bi* 皮痹): numb and cold feeling in the skin

Muscle Bi (*rou bi* 肉痹): numbness and weakness of the muscles, generalise lassitude and easy sweating

Tendon Bi (*jin bi* 筋痹): spasm of the tendons, pain and stiffness of the joints

Vessel Bi (脉痹 *mai bi*): burning sensation of the skin, pain in the muscles and erythematous rash

Bone Bi (*gu bi* 骨痹): pain in the bones, heavy feeling of the body and lameness of the limbs

Legge and Vance (1997) use the terms muscle and tendon Bi along with the standard theory rather than as part of the five phase alternative theory. The signs and symptoms given are not the same as the schema described above. Muscle Bi is identified by muscular pain, stiffness and coldness in muscle groups; tendon Bi by pain radiating along the course of the channel, stiffness, coldness or numbness (Legge and Vance, 1997). Those with knowledge of TBSA may not consider these differences to be irreconcilable. They do, however, represent the typical inconsistencies found in TBSA, which create challenges in establishing the objective disease categories necessary for clinical research.

III.6.2. Japanese acupuncture

The term bi-syndrome does not feature in the Japanese acupuncture textbooks reviewed. The aetiology is sometimes described in terms similar to biomedical

thinking. For example obesity, sports, postural imbalances and the ageing process (Matsumoto and Euler, 2008, Shodo, 2003).

The Japanese textbooks suggest the use of local points (Birch and Ida, 1998, Matsumoto and Euler, 2008, Shodo, 2003). However, the local points are not emphasised to the same degree as they are in the TCM sources. In a case study, Birch and Ida (1998) did not use any local points. Stress is placed on understanding the underlying problem and addressing this before treating the local condition (Matsumoto and Euler, 2008). Distal points are recommended when the knee is acutely painful, with a warning that 'over treatment' could make the situation worse (Shodo, 2003). The suggested local points are similar to the TCM textbooks. In the Japanese texts, palpation is emphasised as a means of finding the points over the classical point location (Birch and Ida, 1998, Matsumoto and Euler, 2008, Shodo, 2003).

The recommended distal points differ from those commonly identified within the TCM textbooks. In particular, the Japanese books recommend using points related to the eight extraordinary vessels (Birch and Ida, 1998, Matsumoto and Euler, 2008). Whilst these points are commonly used in TCM they do not feature prominently in the points recommended for knee pain. Distal points are recommended based on the distribution of pain and meridian pathway, similar to the TCM style (Matsumoto and Euler, 2008).

Point stimulation is different, as Japanese acupuncture often uses shallow insertion (Birch and Ida, 1998). Furthermore, intra-dermal needles are specifically recommended (Birch and Ida, 1998, Matsumoto and Euler, 2008). Additional methods of stimulating points include magnets (Matsumoto and Euler, 2008) and a wooden hammer (Birch and Ida, 1998).

Moxibustion is suggested, with Shodo (2003) agreeing with a Cheng (1999), a TCM source, that moxibustion is sometimes more important than needles. The recommended method is direct moxibustion (Shodo, 2003). Direct moxibustion is also recommended by Legge and Vance (1997), however, in TCM warm

needle is more frequently suggested (Cheng, 1999, Legge and Vance, 1997, Maciocia, 1994, Zhao and Wang, 2007). Warm needle can also be used in Japanese acupuncture (Birch and Ida, 1998).

In common with the TCM textbooks, how long a course of treatment would be expected to last is not clear. Shodo (2003) notes that the condition should be treated early. Palpation around the knee can be used to discover tender points, before an individual develops pain and discomfort in their everyday life (Shodo, 2003). This should not be seen as a specific difference with TCM, but rather the emphasis of an individual author. The concept of early treatment has its origins in the Huangdi Neijing: the expression 上工治未病 *shang gong zhi wei bing* has the meaning of treating before a condition arises. Therefore, even though the TCM textbooks do not highlight early treatment it may well be recommended by many TCM practitioners.

III.6.3. Medical Acupuncture

The medical acupuncture textbooks recommend the use of acupuncture points local to knee (Baldry, 2004, Ma et al., 2005, White et al., 2008). These points should be located by palpation, to find painful and tender areas. White et al suggest that at least one local point should be a traditional acupuncture point (2008). Distal points are generally not emphasised. Although, the lower back should be evaluated and needled when treating disorders of the lower limb (Ma et al., 2005). One distal point is specifically named LR 3 *taichong* (White et al., 2008). Interestingly, LR 3 *taichong* is not cited in any of the TCM textbooks, however, it is a point that maybe used in any painful condition (Deadman et al., 2007).

The needling sensation of deqi is thought by Ma et al (2005) to be important for pain relief. However, they counsel that for certain patients a strong sensation or even a weak sensation can cause a flare up of pain. They recommend starting with an under dose and then increasing stimulation over the course of treatment. Increased stimulation can be accomplished by, using more needles,

manipulating the needles, retaining the needles for longer and the use of electro-acupuncture (White et al., 2008). Shallow needling, while not the standard approach, could also be used (White et al., 2008). Ma et al maintain that experienced practitioners are able to detect a needle grasp sensation. Moreover, that this grasping sensation is related to a 'cascade of downstream physiologic healing effects' (p28, 2005). A minimum of six treatments is recommended (Filshie and White, 1998). Although, in a case study treatment was ongoing every three weeks for four years (Baldry, 2004).

III.6.4. Five Element - Stems and branches

Five-elements and Stems and branches are two separate styles of acupuncture. They are combined in this section because the essential finding of the review is the same for both styles. Treating the whole person is strongly emphasised. The signs and symptoms associated with clinical osteoarthritis do not provide enough information to predict in any way how patients would be diagnosed or treated (Golding, 2013, Hicks et al., 2010, Jarrett, 2006, Jarrett, 1999, Worsley, 1998, Worsley, 2012). In the TCM or Medical acupuncture texts knee pain or osteoarthritis can be located easily through the index. However, this is not the case for Five-element or Stems and branches, reflecting the different theoretical approaches. Five elements considers the spirit (*Shen* 神) to be the fundamental cause of the majority of conditions that present in the West (Hicks et al., 2010). Stems and branches is related to astrology so the patient's date of birth is an essential part of the diagnostic process (Golding, 2013).

III.7. Discussion

The project has two aims, the development of a protocol to investigate warm needle acupuncture and the translational analysis of acupuncture for osteoarthritis of the knee ([section-II.5.4](#)). This review has provided information required for both of these aims.

III.7.1. Limitations

There are some limitations to the narrative review. Firstly, it has not fully covered all the theoretical approaches ([section-III.3](#)). However, there is a high degree of consistency within the TCM textbooks regarding the core pattern differentiation and treatment methods. It is considered highly unlikely that extending the review would shed any more light on the ‘standard theory’, particularly in regards to the treatment of the biao-branch. Where there is heterogeneity, it is found in the assessment and treatment of the underlying-root condition. Further review would doubtless add to this heterogeneity, however, it probably would not change the essential finding. Namely, these differences are most likely due to the personal preferences of the authors or editorial decisions, and do not represent fundamental disagreements regarding the theory.

Secondly, the review was carried out by the researcher who is a TCM practitioner. A practitioner who specialised in another style may have assessed the material differently. Efforts were made to ameliorate this problem by contacting colleagues for advice regarding textbooks and to confirm specific issues. Nevertheless, a collaborative team approach to the review may have produced a more rounded understanding.

Finally, the key limitation of the review is that it only deals with the theoretical approaches. It does not provide direct information regarding acupuncture in practice. Therefore, it is difficult to judge whether the underlying root, the more heterogeneous component of the pattern differentiation, is important in

practice. Similarly, the diversity of the actual practice cannot be assessed. These issues can only be addressed by the expert interviews and practitioner survey.

III.7.2. Assessment of osteoarthritis knee for the pilot study

Based on this review, we can conclude that osteoarthritis of the knee is a suitable condition for the pilot study. A theory common to all the TCM textbooks has been identified, in which there is broad agreement regarding the causes, differential patterns, and treatment strategies for osteoarthritis of the knee. According to the criteria set out in [box-II.2](#), a semi-flexible protocol derived from the standard theory should have a high degree of replicability.

The pilot study ([chapter-VIII](#)) also requires a condition that is suitable for moxibustion. The review has confirmed that moxibustion is an appropriate technique for osteoarthritis of the knee. Moxibustion is recommended by many authors, with some of the opinion that it is more important than needling (Cheng, 1999, Shodo, 2003). The rationale and caveats for selecting osteoarthritis of the knee are summarised in [box-III.4](#).

Box-III.4 Rationale for selection of osteoarthritis of the knee

Warm needle acupuncture is recommended for osteoarthritis of the knee

A semi-flexible protocol should have external validity, because:

- There were relatively few associated differential patterns
- There was a clear overlap in the suggested acupuncture points for both Cold Bi and Damp Bi - which minimises the heterogeneity
- There was general agreement regarding which acupuncture points should be used for each pattern
- Inter-rater reliability should be high for the key signs and symptoms that guide pattern differentiation of the branch

Suitable for moxibustion

The risks to external validity of a semi-flexible protocol

- The relative importance of the root differential pattern
- The extent of diversity in practice

III.7.3. TEAM within a translational research approach

A translational research approach for acupuncture synthesises information from four domains ([section-II.4.6.b](#)). This review provides information on TBSA, that belong to the domain of TEAM. There is no simple chronological relationship between the four domains: the review informs the development of the practitioners survey; the practitioner survey ameliorates the limitations of the review.

The review has provided information on the physical components of acupuncture employed when treating osteoarthritis of the knee. This includes the recommended acupuncture points, methods of needle stimulation, the use of moxibustion and auxiliary techniques. A method of categorising recommended acupuncture points according to the clinical reasoning of the standard theory has been identified ([section-III.5.5.a](#)). This knowledge is used to inform: the interview schedule for the expert interviews ([chapter-V](#)); the development of the practitioner survey questionnaire ([chapter-VI](#)).

A key challenge of acupuncture research is the homogeneity-heterogeneity of praxis ([section-I.3.3](#)). The review has shed some light on this issue, with particular regard to osteoarthritis of the knee, giving some indication of the ways in which styles of acupuncture differ. The differences between TCM and other TBSA echo those seen within TCM. In particular, the heterogeneity tends to be related to the root - underlying condition. Local points around the knee are consistently recommended by Japanese and medical acupuncture textbooks. Japanese acupuncture uses distal points but these are different to those suggested in the TCM textbooks. Medical acupuncture tends not to emphasise the selection of distal points. The Five-element and Stems and branches styles concentrate on treating the root. Moxibustion along with the other techniques proposed by the various authors such as plum needle, cupping and herbs could also be sources of diversity within practice.

The review also indicates areas in which there is synthesis between styles of practice. Within the TCM textbooks, in addition to the standard theory, an alternative theory is described. The alternative theory is unambiguously related to the Five-element style. Ma (2005) a biomedical acupuncture text appears to integrate ideas from TBSA, whereas Baldry (2004) does not. Zhao (2002) and Matsumoto and Euler (2008) use modern anatomical language to a considerable degree.

A narrative review of literature can not assess the extent to which practitioners address the underlying root in practice, or the extent to which auxiliary techniques are employed. Similarly, it cannot establish how commonly dual trained practitioners opt to use the TCM standard theory or take a different approach, such as Five-element or Japanese style. These questions can only be answered using information drawn directly from practice.

IV. Systematic review: warm needle vs.
needle acupuncture

系
统
综
述

IV.1. Introduction

Before conducting a clinical trial it is standard practice to undertake a systematic review, to confirm that a gap in the evidence-base exists. The pilot study protocol is designed to test the hypothesis that warm needle acupuncture improves clinical symptoms more than needle acupuncture, alone ([section-VIII.1](#)). A systematic review of controlled trials comparing warm needle acupuncture to needle acupuncture is presented in this chapter.

Prior to the investigation of the components of a complex intervention, the evidence should indicate that the intervention in its entirety is beneficial ([section-II.3.1](#)). Therefore, a review of systematic reviews with meta-analysis was conducted and the results are presented in this chapter.

The systematic reviews have also analysed acupuncture versus sham comparisons. Whilst often not presented as such in the original papers, sham acupuncture controls are component efficacy investigations ([section-I.4.4](#)). Moreover, when the spectrum of factors known as the placebo effect are balanced between groups, these trials can potentially demonstrate change is due to the physical variables - not the psychologically mediated effects ([section-II.4.4](#)). This project is concerned with the physical components ([section-II.2.2](#)), with the assumption that these components influence clinical outcomes. The validity of this assumption is assessed by evaluating acupuncture versus sham evidence.

The research framework proposes that clinical evidence can only be properly interpreted with reference to the domains of theory and practice ([section-II.4.6.b](#)). The final objective of this review is to gather information on clinical research protocols for the translational analysis of acupuncture for osteoarthritis of the knee ([chapter-VII](#)).

Box-IV.1: Objectives of the literature review

- ▶ Evaluate the evidence for warm needle vs needle acupuncture for osteoarthritis of the knee
- ▶ Evaluate the overall effectiveness of acupuncture for osteoarthritis of the knee
- ▶ Evaluate the evidence for the physical variables of acupuncture influencing clinical outcomes
- ▶ Gather information on clinical research protocols for the translational analysis of acupuncture for osteoarthritis of the knee (chapter-VII)

IV.2. Evidence-base for needle acupuncture

Systematic reviews with meta-analyses are at the apex of the evidence pyramid. Systematic reviews of acupuncture for pain have themselves been reviewed, and the quality was found to be excellent (Lee and Ernst, 2011). Therefore, little would be gained by conducting a full systematic review of systematic reviews, within this project. The objective is to identify relevant systematic reviews, then evaluate whether osteoarthritis of the knee meets the conditions required for this project: evidence of effectiveness ([section-II.5.2](#)); and evidence that the physical variables influence outcomes ([section-II.2.2](#)).

Search strategy

The Medline, CINAHL complete, AMED databases were searched via EBSCO, using the following search terms osteoarth*/arthritis/arthritis/pain AND acupuncture AND systematic review/meta-analysis.

Evaluation of results

A number of systematic reviews of acupuncture for osteoarthritis of the knee have been conducted. The comparisons with no-acupuncture controls showed statistically significant differences favouring acupuncture (Cao et al., 2012, Ezzo et al., 2001, Hou et al., 2015, Lin et al., 2016, Manheimer et al., 2010, Manheimer et al., 2007, Zhang et al., 2017). For example, the comparison with waiting list controls for pain was SMD 0.96 (95% CI, 1.21 to 0.70) (Manheimer et al., 2007).

According to a consensus statement, this should be considered a clinically relevant difference (Dworkin et al., 2008).

The clinical evidence for acupuncture and osteoarthritis of the knee has been included within systematic reviews with wider remits. Reviews of acupuncture and chronic or osteoarthritic pain (Ernst and Lee, 2010, Lee and Ernst, 2011, Madsen et al., 2009, Manyanga et al., 2014, Vickers et al., 2012, Vickers et al., 2017); or reviews on placebo (Hopton and MacPherson, 2010, Hróbjartsson and Gøtzsche, 2010); and reviews focused on a condition with various physical therapies, that included acupuncture (Bjordal et al., 2007, Corbett et al., 2013).

There is considerable overlap with the same clinical trials being evaluated in these different systematic reviews. The reviews consistently conclude that acupuncture is superior to non-acupuncture controls. However, the quality of the included studies was often considered variable (Lee and Ernst, 2011). A landmark review published in 2012 included only high-quality RCTs (Vickers et al., 2012). The meta-analysis within this review uses raw patient rather than summary data, and so has greater statistical precision. Vickers et (2012) demonstrate that acupuncture is effective compared to no-acupuncture controls (SMD -0.57, 95% CI -0.64 to -0.50). *The first pre-condition for conducting a component efficacy trial has been met: acupuncture in its entirety is effective for osteoarthritis of the knee.*

The reviews of osteoarthritis of the knee present consistent results, in favour of acupuncture, for the more contentious comparison with sham acupuncture (table-IV.1) There are, however, slight differences in tone regarding the interpretation of these results. From 'clinically irrelevant short-term improvements' (Manheimer et al., 2007, p686), to 'these benefits are small, do not meet our pre-defined thresholds for clinical relevance' (Manheimer et al., 2010, p2), to 'the effect size compared to sham is 0.4 which is considered "moderate"' (White et al., 2007, p389). Whilst these are all more positive than

interpretations given elsewhere (section-I.4.2), they illustrate that the threshold for a clinically relevant effect size is a subjective judgement.

Table-IV.1: Short term pain outcome compared to ‘sham’ acupuncture

Review	SMD	95% CI
Manheimer et al 2007	-0.35	-0.55 to -0.15
White et al 2007	-0.4	-0.6 to -0.1
Manheimer et al 2010	-0.29	-0.49 to -0.08
Cao et al 2012	-0.25	-0.42 to -0.09

SMD - standardised mean difference: CI- Confidence Interval

Among the reviews with wider remits, Madsen et al (2009) focus on the comparison with placebo and conclude ‘a small analgesic effect of acupuncture was found, which seems to lack clinical relevance and cannot be clearly distinguished from bias’ (p1) (SMD -1.7, CI -0.26 to -0.08). Madsen et al (2009) combine the results of various conditions, including osteoarthritis, migraine, fibromyalgia, scar pain, colonoscopy. However, this is a generalised result that does not consider that some conditions may be more likely to respond to acupuncture others. Manyanga et al (2014) draw a more positive conclusion, ‘acupuncture use was associated with significant reductions in pain intensity (SMD -0.29, CI -0.55 to -0.02)’ (p1). This statistic combines results from different control groups: sham, usual care, and waiting list.

Regardless of the remit of the review or the interpretations of the authors, the sham procedures are consistently shown to be statistically inferior to acupuncture. It is not the pure mathematical analysis of the data that differs, rather value-based judgements in framing of the research and interpretation (section-II.2). As the psychologically mediated effects are balanced between groups in sham controlled trials (section-I.4.4), *the evidence demonstrates that the physical components of acupuncture play a role in the overall effectiveness.* Because of the heterogeneity within these reviews, the conditions included and

the controls used, assessing the effects size as being clinically relevant, or not, is of limited value.

The analysis by Vickers et al demonstrates 'a robust difference between acupuncture and sham control that can be distinguished from bias' (2012, E7) (SMD -0.26, 95% CI -0.34 to -0.17). This meta-analysis included only high-quality trials, did not combine results for disparate conditions or various controls, and has greater statistical precision than other reviews. Yet, the extent to which meaningful inferences can be drawn regarding the size of the effect remains limited. This is because there are two important issues that Vickers et al (2012), and other systematic reviews, have not fully addressed; namely, an assessment of the external validity of the acupuncture procedures and the risk of bias from the sham protocols ([section-II.5.4](#)). The translational analysis ([chapter-VII](#)) evaluates the acupuncture protocols and sham procedures of the trials included in the update of Vickers et al (2012) (Vickers et al., 2017). This is a novel method of evaluating external validity and risk of bias; it is based on data drawn from the practitioner survey ([chapter-VI](#)).

IV.3. Methods - Systematic review

The primary objective of this systematic review is to evaluate the clinical evidence for warm needle acupuncture versus needle acupuncture. The secondary objective is to gather data on the acupuncture protocols for the translational analysis ([chapter-VII](#)).

IV.3.1. Study Criteria

Types of studies

Controlled-trials comparing warm needle acupuncture to needle acupuncture written in English or Chinese were included. Trials published in other languages were excluded, as no funds were available for translation. A preliminary search indicated that the reporting and trial quality was low. Therefore, no restrictions were placed on the type of outcome measures or on

the duration of the observation period. This would enable key weaknesses in the current evidence to be identified, and allow the second objective of mapping and evaluating the acupuncture protocols to be achieved.

Participants

Only trials that dealt directly with osteoarthritis of the knee were included, trials that also addressed hip/back pain in addition to the knee were excluded. No restrictions were placed on the diagnostic method used in the inclusion criteria.

Types of intervention

Studies were included only if, they compared warm needle acupuncture to needle acupuncture. Studies were excluded if, they employed other methods of moxibustion such as moxa stick or scarring moxibustion.

Trials were excluded if they adopted an integrative approach, by providing additional therapies (such as massage, Chinese herbal medicine or corticosteroid injections) to one or both groups. Studies were included if the only other intervention was a standard pharmaceutical medication, such as non-steroidal anti-inflammatory drugs (NSAID), provided to both the experimental and control groups. Protocols that used electro-acupuncture or electronic moxa machines were excluded. Studies were included regardless of the acupuncture theory being used to make the acupuncture points selection; Traditional Chinese medicine, five element or any other theory could be included.

Outcome measures

No restrictions were placed on the type of outcome measure used.

IV.3.2. Search strategy

Searches were carried out in both English and Chinese. The Chinese Knowledge Infrastructure (CNKI) database was used to access studies published in China. [Box-IV.2](#) shows the Chinese language search strategy. A discussion of the particular challenges of the Chinese language search is presented in [section-IV.6](#).

Box-IV.2: Chinese language search strategy

SU =(温针灸+温针+温针刺+温针治疗) AND SU = (膝痹+膝痛+膝关节) AND
FT = (随机+对照) NOT TI = (注射+电针+玻璃+中药+推拿)

A parallel search was conducted using EBSCOhost to access Medline, Cumulative Index to Nursing and Allied Health Literature (Cinahl Complete) and the Allied and complementary medicine database (AMED). The search terms are set out in [box-IV.3](#).

Box-IV.3: English language search terms

- acupunct*
- knee
- moxa OR moxibustion OR moxabustion
- pain OR arthritis OR osteoarth* OR gonarthrititis

IV.3.3. Data extraction and analysis

Analysis of the retrieved studies was carried out using the standard categories of: Participants, Intervention, Control and Outcome measures (PICO). The Cochrane Handbook was used as a guide to assess the Risk of bias (Higgins and Green, 2008). The CONSORT extension for acupuncture, items 2-4 (MacPherson et al., 2010), was used as a guide to identify the components of acupuncture.

Ideally, evaluation of the risk of bias would have been carried out by two or more evaluators who were blinded to the names of the authors and the journal of publication (Higgins and Green, 2008). This research was conducted as part

of a PhD project and, consequently, it was not possible for each research paper to be evaluated by two independent researchers or blind the assessor.

As a PhD student the assessor was, almost by definition, a novice researcher. Therefore, a preliminary evaluation of the English language evidence for acupuncture treating osteoarthritis of the knee was conducted first. The researcher was able to compare his findings to those published in the Cochrane systematic review of acupuncture for peripheral joint osteoarthritis (Manheimer et al., 2010). The evaluation of certain aspects of the papers were discussed with supervisors.

The researcher is a second language speaker of Chinese. When there were areas of ambiguity regarding language issues, advice was sought from colleagues who were native Chinese language speakers working at London South Bank University.

IV.3.4. Risk of bias assessment

The Cochrane handbook was used to assess the risk of bias (Higgins and Green, 2008). The handbook sets out six criteria for judging the risk of bias: selection bias, performance bias, detection bias, attrition bias, reporting bias and other bias. For each item the judgement should be categorised as high risk, low risk or unclear risk.

Information supporting each judgement is presented in [appendix-C.1](#). When the judgement was made on a specific sentence, the original Chinese has also been presented in the table ([appendix-C1](#)). Within the *other bias section* the Outcome Measure used was frequently identified as a potential risk. The outcome measures are discussed in detail in [section-IV.4.4](#).

IV.3.5. Quantitative data analysis

Due to the poor quality of the trials retrieved (see below) no quantitative data analysis was undertaken. The reported results were compiled into tables to enable descriptive discussion of the current evidence base.

IV.4. Results - systematic review

IV.4.1. Included Studies

All of the retrieved trials were conducted in China and published in the Chinese language. A total of 486 results were returned, 42 studies were retrieved for detailed examination. One study was a duplicate, retrieved via both EBSCOhost and CNKI (Xiong et al., 2011). There were 34 trials that met the criteria (figure-IV.1). The majority of these studies were non-blinded controlled trials that compared warm needle acupuncture to needle acupuncture (Cui, 2013, Hu et al., 2014, Huang, 2007, Li et al., 2006, Li et al., 2011, Li and Li, 2010, Liu et al., 2003, Luo and Mao, 2011, Lv, 2000, Wang, 2007, Wang and Zhu, 2009, Xia, 2011, Yan et al., 2013, Yang, 2014, Yao et al., 2003, Yu, 2009, Zhang, 2013a, Zhang, 2013b, Zhao, 2014, Zhao et al., 2008). Two studies used a TDP heat lamp as a method of masking the radiated heat of the warm needle acupuncture, and blind the participants to group allocation (Yue, 2010, Zhang et al., 2009).

Excluded Studies

Seven studies were excluded from the review. Two studies used different acupuncture points between the experimental and control groups (Fan, 2014, Ma, 2016). One study employed an electro-moxa machine (Zhu et al., 2011), another used a heat lamp (Xiong et al., 2011). A study by Yu Jian (Yu, 2008) was found to be a duplicate of Yu Jian (2009) - despite the two papers having different titles. The 2009 publication was included in the review. Liu and Du (2015) was actually a review paper. Li 2010 was a duplicate - (Li et al., 2011).

In most studies the participants were hospital outpatients. Some studies did not report the setting explicitly but it can be inferred from statements such as "my department". The studies included participants of broadly similar age ranges. There appears to be heterogeneity in the duration of knee pain. Over half of the studies did not report a mean for the duration of the knee pain, the ranges were more frequently reported. One study had a range of 3 to 15 years (Xia, 2011) whilst another reported a range of only 1 week to 3 years (Yao et al., 2003).

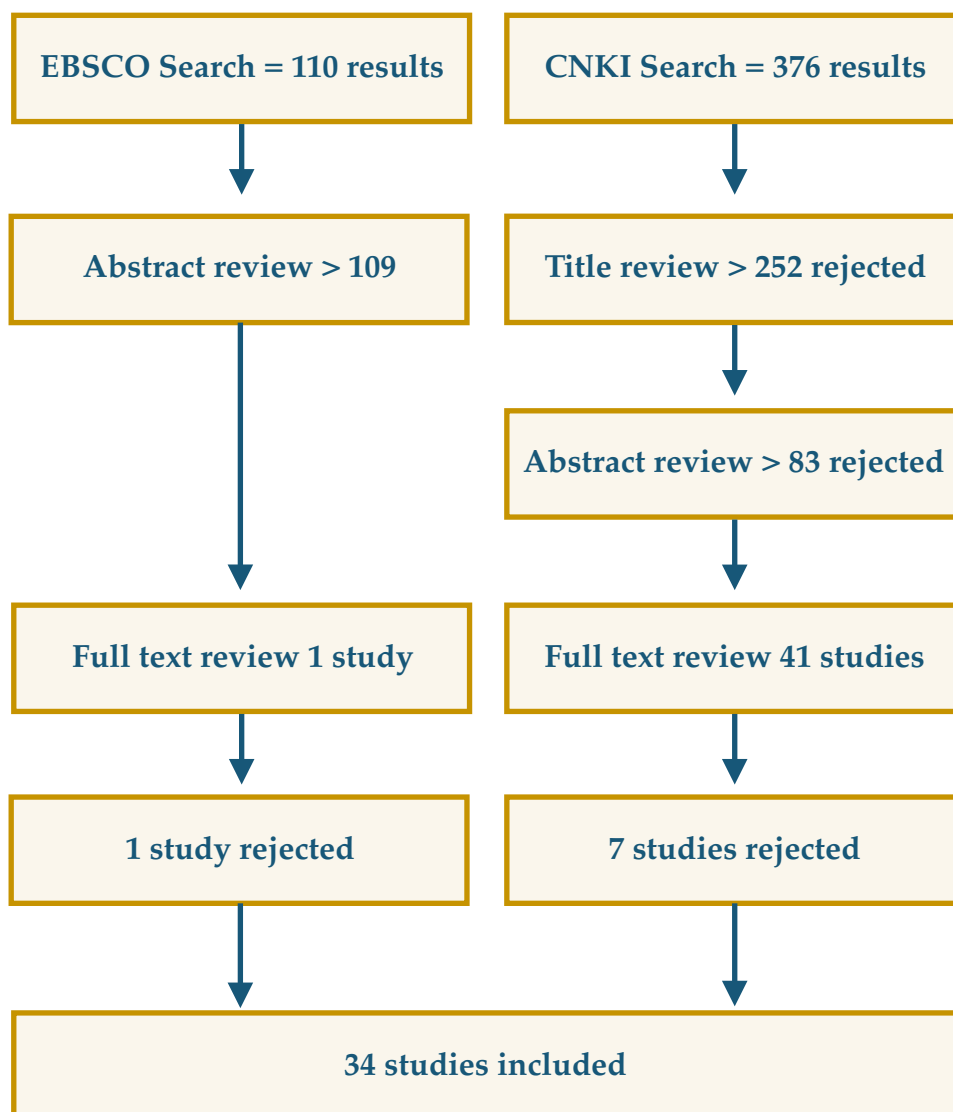
The majority of studies utilised the American College of Rheumatology diagnostic criteria as part of the inclusion criteria. Some studies referred to Chinese standards (Yan et al., 2013, Yang, 2014, Zhao, 2014), others reported the signs and symptoms without a clearly referenced criteria (Cui, 2013, Li and Li, 2010), others simply stated that the participants had osteoarthritis of the knee (Zhang, 2013a, Zhang, 2013b). Only four studies reported baseline statistics based upon a quantitative scale, to measure pain and/or disability (Li et al., 2011, Li and Li, 2010, Zhang, 2013a, Zhang et al., 2009).

IV.4.2. General characteristics

A total of 3134 participants were involved in the studies identified in this review. The largest study had 300 participants (Zhang et al., 2009) the smallest 58 (Yang, 2016). The main characteristics of the studies are presented in [appendices-C.2-5](#).

The generally poor standards in reporting, in particular the lack of information regarding baseline measures, meant it is not possible to judge whether the participants were broadly similar prior to treatment. Although, they appeared to be approximately similar in age, 40-75 years old. A variety of different outcome measures were employed, these are discussed below. None of the studies reported funding sources.

Figure-IV.1: Study selection flowchart



IV.4.3. Risk of bias of included studies

The risk of bias for all 34 trials is unclear ([table.IV.2](#)), mostly due to poor standards of reporting. The risk of bias assessments for individual studies are presented in [appendix-C.1](#). The majority of trials report that participants were randomised. Many, 17/34, simply report that randomisation was used for group allocation and no further information is given regarding the methods used; such as (Cui, 2013, Wang and Zhu, 2009). These have been classed as having an unclear risk of bias. Randomisation using an appropriate method was reported by 11/34 trials, including: computer generation (Ding, 2016) and the drawing of lots (Liu et al., 2003); the most commonly used method was a random numbers table, e.g. (Tan and Feng, 2016, Yan et al., 2013). These have been assessed as low risk of bias. Four trials used the order in which participants sought medical help or their patient record numbers as a means of randomisation; these have been classed as high risk (Hu, 2015, Liu, 2016, Wu, 2015, Xia, 2011). Two trials that provided no information at all have been classed as high risk (Lv, 2000, Zhao et al., 2008).

None of the trials report any method of allocation concealment, the two non-randomised trials have been classed as high risk (Lv, 2000, Zhao et al., 2008). The other 32 trials have been assessed as having unclear risk of bias for allocation concealment.

Performance bias refers to the methods of blinding the participants and practitioners. The majority of studies do not describe any methods for blinding participants or practitioners, 32/34. The physical nature of acupuncture means any blinding would require specific procedures to be developed. Therefore, in the absence of any description of blinding procedures it has been assumed that no attempt of blinding was made. These trials have been assessed as high risk. Only Yue (2010) and Zhang et al (2009) attempted to blind the participants. In both cases this was done using a heat lamp, to mask the radiated heat of moxibustion. Whether these procedures were successful, or not, is not reported. Nor is there any discussion regarding the blinding of those administering the

acupuncture (Yue, 2010, Zhang et al., 2009). Both trials have been assessed as having an unclear risk.

Detection bias stems from those charged with collecting outcome data being aware of group allocation. None of the trials report how the data was collected, therefore, the risk of bias is unclear.

Attrition bias refers to the influence of participants not completing a study may have on the results. The majority of studies 32/34 appear to include all participants in the final analysis, however, none discuss the lack of attrition in their respective studies. Only two studies reported attrition (Li et al., 2006, Yan et al., 2013), both have been classed as a low risk of bias. For all the other studies the risk is unclear. None of the studies report pre-publication protocols, therefore, all have been classed as having an unclear risk of *reporting bias*.

Table-IV.2: Risk of bias summary - researchers judgement

	Random sequence	Allocation concealment	Blinding of participants/ personnel	Blinding of outcome assessment	Incomplete outcome data	Selective reporting	Other
High risk							
Unclear risk							
Low risk							
Cui Hw 2013							
Ding Jx 2016							
Hu Jh 2015							
Huang Cx 2007							
Li Cd 2006							
Li L 2011							
Li Zh 2010							
Liu B 2016							
Liu La 2003							
Lu Xh 2015							
Luo Xf 2011							
Lv M 2000							
Qin D 2016							
Tan Hy 2016							
Wang Gz 2007							
Wang Gz 2009							
Wu Hm 2015							

Table-IV.2: Risk of bias summary - researchers judgement (continued)

High risk							
Unclear risk							
Low risk							
	Random sequence	Allocation concealment	Blinding of participants/ personnel	Blinding of outcome assessment	Incomplete outcome data	Selective reporting	Other
<u>Xia Qf 2011</u>	High risk	Unclear risk	High risk	Unclear risk	Unclear risk	Unclear risk	Low risk
<u>Xu GY 2016</u>	Unclear risk	Unclear risk	High risk	Unclear risk	Unclear risk	Unclear risk	Unclear risk
<u>Yan Xl 2013</u>	Low risk	Unclear risk	High risk	Unclear risk	Low risk	Unclear risk	Unclear risk
<u>Yang D 2014</u>	Low risk	Unclear risk	High risk	Unclear risk	Unclear risk	Unclear risk	High risk
<u>Yang Ky 2015</u>	Low risk	Unclear risk	High risk	Unclear risk	Unclear risk	Unclear risk	Low risk
<u>Yang Xs 2016</u>	Low risk	Unclear risk	High risk	Unclear risk	Unclear risk	Unclear risk	Unclear risk
<u>Yang Ys 2016</u>	Unclear risk	Unclear risk	High risk	Unclear risk	Unclear risk	Unclear risk	High risk
<u>Yao Zf 2003</u>	Unclear risk	Unclear risk	High risk	Unclear risk	Unclear risk	Unclear risk	High risk
<u>Yu B 2015</u>	Unclear risk	Unclear risk	High risk	Unclear risk	Unclear risk	Unclear risk	Unclear risk
<u>Yu J 2009</u>	Unclear risk	Unclear risk	High risk	Unclear risk	Unclear risk	Unclear risk	High risk
<u>Yue R 2010</u>	Unclear risk	Unclear risk	Unclear risk	Unclear risk	Unclear risk	Unclear risk	High risk
<u>Zhang Gc 2015</u>	Unclear risk	Unclear risk	High risk	Unclear risk	Unclear risk	Unclear risk	Unclear risk
<u>Zhang J 2013</u>	Low risk	Unclear risk	High risk	Unclear risk	Unclear risk	Unclear risk	Low risk
<u>Zhang Jf 2009</u>	Unclear risk	Unclear risk	Unclear risk	Unclear risk	Unclear risk	Unclear risk	Unclear risk
<u>Zhang Zp 2013</u>	Unclear risk	Unclear risk	High risk	Unclear risk	Unclear risk	Unclear risk	High risk
<u>Zhao M 2014</u>	Unclear risk	Unclear risk	High risk	Unclear risk	Unclear risk	Unclear risk	High risk
<u>Zhao Xt 2008</u>	High risk	High risk	High risk	Unclear risk	Unclear risk	Unclear risk	High risk

IV.4.4. Other Bias - Outcome measures

In the majority of studies, the outcome measures appear to be unvalidated instruments and constitute a risk of bias. The study is present their results using a variety of formats, to facilitate analysis these have been organised into three groups: original categories, scale responder- rates and scale.

A number of studies use more than one outcome measure, which have different levels of risk. [Table-IV.2](#) shows the judgement on the outcome measure with the lowest level of risk. In others words, if a study used two measures one judged to be high risk the other as unclear, then the study is identified as having unclear risk in [table-IV.2](#). This allows a degree of differentiation between studies to be illustrated. An overview of the outcome measures used in each study is presented in [appendix-C.6](#).

IV.4.4.a. Ordinal categories as outcome measures

This was the most commonly used method of assessing the therapeutic outcomes. Essentially, basic symptomatic descriptions are given and the patient's knee pain is assessed as belonging to one of 3, sometimes 4, ordinal categories. For example, "The pain has disappeared. Joint functions normally on flexion and extension. There is no relapse within three months" (Lv, 2000); or "The knee pain, swelling and other symptoms have been alleviated. The function of the knee has improved" (Yang, 2016); or "There is no obvious change in the pain and stiffness of the affected joint. There is no improvement in the movement and function" (Huang, 2007). See [appendix-C.7](#) for full table, abridged version [table-IV.4](#)

A number of terms are used to describe the categories: ranging from "Completely cured" *Quányù* 痊愈 to the optimistic "Yet to recover" *Wèi yù* 未愈. To allow comparisons between trials these terms have been converted to a 4 level numeric system; 0 is the least effective category and 3 the most effective. The original category names and the corresponding levels are set out in [table-IV.3](#). "Markedly effective" is usually the second level of improvement.

However, some authors use this term for most improved situation (Luo and Mao, 2011), in these cases “Markedly effective” has been classed as level 3.

Table-IV.3: Category Names

Chinese	Pinyin	Translation	Level
痊愈	<i>Quányù</i>	Completely recovered	3
治愈	<i>Zhìyù</i>	Cured	
优	<i>Yōu</i>	Excellent	
基本治愈	<i>Jīběn zhìyù</i>	Basic cure	
临床治愈	<i>Línchuáng zhìyù</i>	Clinically cure	
临床痊愈	<i>Línchuáng quányù</i>	Clinically completely recovered	
临床控制	<i>Línchuáng kòngzhì</i>	Clinically controlled	2
显效	<i>Xiǎnxiào</i>	Markedly effective	
有效	<i>Yǒuxiào</i>	Effective	1
好转	<i>Hǎozhuǎn</i>	Improvement	
良	<i>Liáng</i>	Good	
可	<i>Kě</i>	Possible	0
无效	<i>Wúxiào</i>	Ineffective	
差	<i>Chà</i>	Poor	
未愈	<i>Wèi yù</i>	Yet to recover	

Overall the categories used in different studies appear broadly similar, however, this is primarily due to the categories themselves being vague. None of the trials report how the assessments were made, or by whom. There must have been cases which did not fit cleanly into one of the categories. No explanation is given regarding the difference between pain being reduced or significantly reduced. Nor is there an indication of what happened when the swelling reduced but the pain did not, or vice versa. Another problem with this method of evaluation is the lack of baseline data. It is not possible to ascertain whether the degree of pain and stiffness were broadly similar between groups at baseline. Consequently, all studies that used ordinal categories outcome measures as a sole means of evaluation have been classed as high risk of bias; 13 trials (Huang, 2007, Li and Li, 2010, Lu, 2015, Luo and Mao, 2011, Lv, 2000, Wu, 2015, Yao et al., 2003, Yu, 2009, Yue, 2010, Zhang et al., 2009, Zhang, 2013b, Zhao, 2014, Zhao et al., 2008).

Table-IV.4: Categories outcome measures (abridged)

	3	2	1	0
Lu Xh 2015		Joint pain and swelling disappear, normal function of the knee joint recovered	The joint pain is relieved or disappears, the swelling has significantly subsided, ability to flex and extend has been recovered	The joint symptoms have not improved and may even have got worse
Luo Xf 2011		The pain and stiffness of the affected knee has disappeared. When moving the joint the sound of crepitus has disappeared or only occasionally seen. Normal movement and function has been recovered.	The pain has reduced, there is still crepitus. Joint function has clearly improved	There is no obvious change in the pain and stiffness of the affected joint. There is no improvement in the movement and function.
Lv M 2000	The pain has disappeared. Joint functions normally on flexion and extension. There is no relapse within three months.	The pain has disappeared there is no obvious limitation of movement in the joint. There is occasional loss of strength when walking.	There is a clear reduction in pain sometimes there is limitation in the movement of the joint. After too much movement the pain returns.	After two courses of treatment the person does not feel there has been any change in the signs and symptoms when compared to pre-treatment
Qin D 2016	The knee pain and swelling have disappeared, the joint function has returned to normal. An X-ray scan reveals no bony spurs.	The knee pain and joint swelling have been significantly alleviated, the joint function has returned to normal. An x-ray scan reveals no bony spurs.	The knee pain and joint swelling have been alleviated, there is also joint function recovery. An x-ray scan reveals no obvious bony spurs.	The knee pain and swelling have not improved, the joint function has not recovered. Bony spurs are revealed after an x-ray scan.

IV.4.4.b. *Scale responder-rates outcome measures*

The second method is the use of a scale with the results converted to responder-rates. The categories of the responder rates use the same terminology employed for the ordinal category method [table-IV.3](#). Participants are identified as being “improved” or “markedly effective” based on the degree of change in the scale. A total of 11 trials use this method; (Cui, 2013, Ding, 2016, Hu, 2015, Li et al., 2006, Li et al., 2011, Liu et al., 2003, Wang, 2007, Wang and Zhu, 2009, Xia, 2011, Yan et al., 2013, Yang, 2014). See [appendix-C.8](#) for full table, abridged version [table-IV.5](#).

A number of different scales were used. In most cases minimal information is reported. Three of the trials describe changes in points scored: Xia, (2011), use the Lysholm Index where ‘good’ is a change of between 11 and 29 points; Liu et al (2003) use a point-score instrument developed in Japan. Both of these trials have been judged as low risk. Li et al (2011) use their own point system. In this system the best outcome is a score of 0-1 points, but the ‘markedly effective’ category is a proportional change of more than two thirds. Because of the inherent inconsistency within their system Li et al (2011) has been judged to have a high-risk of bias.

The majority of trials define the responder-rates by percentage change in the outcome measure: $[(\text{pre-treatment score} - \text{post-treatment score}) / \text{pre-treatment score}] \times 100$. This calculation is often described as 尼莫地平法 *nimodipingfa*, the Nimodipine method in Chinese. Although the reason for naming it as such is unclear. Nimodipine is a calcium channel blocker (NICE, 2014a). Unfortunately, trials use different ranges of percentage change for the ordinal categories. For example, the level 2 category could be 50-80% (Cui, 2013, Li et al., 2006), 60-90 % (Ding, 2016, Hu, 2015), or 70-95% (Yan et al., 2013, Yang, 2014), ([table-IV.5](#); [appendix-C.8](#)). This makes comparison of results difficult.

There is a general lack of clarity regarding the actual scales used. Minimal information is provided in the trial reports (Cui, 2013, Ding, 2016, Hu, 2015, Li et al., 2006, Li et al., 2011, Liu et al., 2003, Yang, 2014). This problem is compounded by the citations. For example, Cui (2013) cite another trial Zhang et al (2009). Zhang et al (2009) describe three different outcome measures but not the scale responder-rates that Cui (2013) used. Therefore, it is not clear how Cui (2013) measured change. Hu (2015) and Liu (2016) also refer readers to other clinical trials.

Three trials cite 'Guiding principles of clinical research of new medicine in Chinese medicine' by Zheng (2002) as the source of their outcome measure: Li Cd et al (2006), Yang (2014) and Li et al (2011). However, Li Cd et al (2006) and Yang (2014) use different percentage ranges. Whilst Li et al (2011) do not use a percentage change but a point-score system instead. Zheng (2002, p351) actually presents a table with suggestions as to how to grade various signs and symptoms as being mild (轻-*qing*), moderate (中-*zhong*) or severe (重-*zhong*). For example, the ability to climb a standard staircase should be assessed as follows: able = mild, with difficulty = moderate, unable = severe. No direct instructions are provided as to how these grades should then be quantified.

Zheng (2002, p352) also suggest the therapeutic effect should be assessed using four grades: clinically controlled, markedly effective, effective, and no effect. For each of these grades changes in signs and symptoms, reduction in point score and X-ray assessment are all described.

For example, markedly effective:

- Pain and other signs and symptoms disappear, there is no limitation in the joint function
- The points score is reduced by more than 70% but less than 95%
- X-rays show a clear improvement

There is ambiguity regarding whether symptoms, the pain score, or x-rays are the defining criteria for assessing the clinical outcome. Consequently, if an

ambiguous scale is the only outcome measure reported the trials have been assessed as high risk (Cui, 2013, Li et al., 2011, Yang, 2014).

Wang and Zhu (2009) and Wang⁹ (2007) appear to be missing a responder-rate category between 70-30%. These trials had 67 and 96 participants, respectively. There is no discussion as to why none of the participants had a reduction of between 70 to 30%. Both of these trials have been classed as high risk.

Yan et al (2013) state that they used WOMAC, however, the citation is for another study not the developer of WOMAC. In addition, the description of categories not only includes changes in the score but also brief comments about the signs and symptoms; classed as unclear risk.

Xia (2011) used the Lysholm index, explain the scoring system and cite Lysholm and Gillquist (1982). Liu et al (2003) used a points-based outcome measure developed in Japan. There is a clear citation. These outcome measures have been assessed as low risk.

⁹ Wang Gz is the author of both studies

Table-IV.5: Scale responder rates (abridged)

	3	2	1	0
Cui Hw 2013	Fundamentally cured: condition severity index greater than 80%	Marked effect: condition severity index greater than 50%	Effective: condition severity index greater than 25%	No effect: condition severity index less than 25%
Ding Jx 2016	Effectiveness index $\geq 90\%$	Effectiveness index $\leq 60\% < 90\%$	$20\% \leq \text{effectiveness index} < 60\%$	Effectiveness index $< 20\%$
Xia Qf 2011*	The difference between the before and after treatment scores ≥ 30	The difference between the before and after treatment scores 11-29	The difference between the before and after treatment scores 6-10	The difference between the before and after treatment scores < 5
Yan Xi 2013	Cured: Signs and symptoms have disappeared or fundamentally disappeared, the index has reduced by $\geq 95\%$	Markedly effective: signs and symptoms have obviously improved, the index has reduced by more than $\geq 70\% < 95\%$	Effective: the signs and symptoms have improved, the index has reduced by $\geq 30\% < 70\%$	Ineffective: no obvious improvement in signs and symptoms, they may have even got worse, the index has not reduced by 30%
Yang D 2014	The reduction in the symptom score $\geq 95\%$	The reduction in the symptom score $\geq 70\% < 95\%$	The reduction in the symptom score $\geq 30\% < 70\%$	Reduction in the symptom score did not reach 30%

* used Lysholm

IV.4.4.c. *Scale outcome measures*

Twelve studies report changes in mean scores ([appendix-C.6](#)). A number of different scales were used: Lequesne, Lysholm Index, WOMAC, Visual Analogue Scale (VAS) and HSS (Hospital for special surgery)¹⁰. In some cases the scale is not clearly described.

Three trials used both the Lysholm Index and a VAS (Ding, 2016, Hu, 2015, Liu, 2016). Ding (2016), Hu (2015), along with another study Xu (2016), used a VAS to measure pain, but in each case the exact question participants were given is unclear. Ding (2016) state the Lysholm Index measures joint function, whereas for Hu (2015) it evaluates knee ligament damage. Liu (2016) state they evaluated quality-of-life, but report no details. None of these studies provide references to indicate where they had accessed the Lysholm Index. These studies have been assessed as unclear risk. Two studies used a HSS index, both have been assessed as unclear risk as neither report details or a clear citation (Yang, 2016, Zhang, 2015).

Li Cd et al (2006) used Lequesne citing a Chinese language textbook as the source. A condition severity index and pain scores are reported. They also utilise a scale developed in China to evaluate the change in pattern differentiation ([section-I.3.5](#)). This study's outcome measure has been assessed as low risk.

Tan and Feng (2016) and Yang and Qu (2015) used VAS to assess pain whilst walking on the flat, assessed as low risk. In addition, both studies assessed lower limb function. Tan and Feng (2016) provide no information regarding this outcome measure other than it was developed in Japan. Yang and Qu (2015) report incomplete details. Neither of these studies provide a citation. These outcome measures have been assessed as unclear risk.

Zhang, J (2013a) used WOMAC; low risk. Zhang et al (2009) state that they used both WOMAC and Lequesne, but provide no citation related to the developers

¹⁰ HSS was assumed to mean Hospital for Special Surgery - the abbreviation was not defined in the trial reports

of these outcome measures. The results are reported in tables labelled 'pain' and 'knee function'. It is not clear whether the two outcome measures were combined, or 'pain' refers to WOMAC and 'knee function' to Lequesne. Zhang et al (2009) has been assessed as unclear risk.

IV.4.4.d. Risk of bias summary

There are two principal problems, reporting standards are low and outcome measures appear to be unreliable. None of the studies provide adequate information. In some cases the description of research methods is so minimal as to render the report little more than an abstract. Of particular concern is the lack of sufficient information on how the data were collected, because the outcome measures themselves often appear to have a high risk of bias. In addition, no studies report funding sources or conflicts of interest. The risk of bias is either high or unclear for all the included studies. Therefore, no firm conclusions can be drawn regarding the additional benefits of warm needle acupuncture.

IV.4.5. Participants

The participants were approximately the same age in all the studies, typically between 40 and 75 years old. Around half of the studies used the American Rheumatoid Association (ARA) or American College of Rheumatology (ACR) standards for the clinical diagnosis. Several studies used Chinese National criteria. Some simply describe signs and symptoms and others provide no information. Half of the trials indicated the use of radiographic diagnoses. However, only Li, Cd et al (2006) describes a minimum pain score, and just 2 studies clearly state a minimum duration as entry criteria (Hu, 2015, Qin, 2016) ([appendix-C.3](#)).

The majority of studies report the means and ranges of illness duration, 7 provide no information. The mean illness duration ranges from 2 months (Xu, 2016) to 3.5 years (Hu, 2015). Only 12 studies report baseline pain scores. All the studies report that the experimental and control groups were comparable at baseline. It is difficult to assess whether the severity the clinical symptoms was

comparable between studies as the outcome measures were different, and often unreliable (section-IV.4.4).

IV.5. Analysis of outcomes

All the studies report positive results in favour of warm needle acupuncture. The results for original categories and scale responder-rates are presented in the appendices-C.9/10. Four studies report mean change in a pain scale (table-IV. 6-8). Zhang J (2013) report the outcome at three months follow-up, whereas the other trials provide end of treatment scores (Li et al., 2006, Tan and Feng, 2016, Yang and Qu, 2015). These studies found statistically significant differences between the warm needle group and needle acupuncture controls. A meta-analysis was not conducted because of the uncertainty regarding the risk of bias (section-IV.4.4). The results are broadly similar.

Table-IV.6: Li Cd et al 2006 - Lequesne pain score change

Group	n	Baseline	after 1 course	after 2 course
Warm	56	6.94 ±1.17	4.71 ± 1.11*	2.81 ±0.92*
Needle	54	7.13 ±0.98	5.66 ±1.23	4.21 ±0.98

* note - between group comparison $p < 0.05$

Table-IV.7: Zhang J 2013 - WOMAC pain score change

Group	n	Baseline	3 months
Warm	45	60.07 ±6.54	26.45 ±5.21*
Needle	45	59.84 ±6.14	40.58 ±4.91

* note - between group comparison $p < 0.05$

Table-IV.8: VAS flat walking test pain

	Group	n	Baseline	Post treatment
Tan & Feng 2016	Warm	40	68.57 ± 19.12	27.52 ± 12.16 [*]
	Needle	40	70.13 ± 20.05	41.61 ± 21.97
Yang & Qu 2015	Warm	68	68.26 ± 14.26	21.22 ± 15.36 [*]
	Needle	68	67.25 ± 15.02	42.39 ± 13.28

* note - between group comparison $p < 0.05$

IV.6. Searching Chinese language databases

The China Knowledge Resource Integrated Database (CNKI) was used to access research published in China. The methods used to search for clinical trials published in the Chinese language are set out in detail in this section. A specific problem related to the Chinese language was identified. Awareness of this issue will be beneficial to those seeking to systematically identify research articles, in the future.

CNKI 'is the most comprehensive knowledge system that integrates 90% of China knowledge and information resources' (CNKI, 2017). The database has a number of different search options including *search*, *advanced search* and *specialist search*. Advanced search and specialist search were both used in the course of conducting this literature review.

The advanced search function was first used and this provided a preliminary understanding of the research available. The advanced search function requires individual search terms to be entered into text boxes. Drop-down menus allow search field conditions to be placed on these search terms, for example 'keyword' or 'full text', as well as Boolean operations AND/ NOT/ OR. Initial searches using this method recovered 11 relevant papers.

The author is not a native speaker of Chinese and previously had minimal experience of using Chinese language databases. Advice was sought from two

colleagues, both of whom are native speakers of Chinese with experience of using the CNKI database. One had completed a PhD in the UK that required searching Chinese language databases, the other was a PhD student at the time. Both colleagues recommended using the specialist search function. The specialist search function consists of one dialogue box into which the entire Boolean search algorithm is entered. A search strategy was developed in collaboration with these colleagues.

There are three components to the CNKI specialist search function. The first is the search fields ([table-IV.9](#)). The second is the use of mathematical symbols to refine searches within a specific search field. For example the + sign operates as an OR function, * sign operates as an AND, - sign operates as a NOT. There are other ways in which the search can be refined, for example by stipulating how many times a term needs to appear within the text or that two terms should appear within the same sentence. The search fields can then be combined using the third component the OR, AND, NOT functions.

An initial search using the specialist search function returned over 2800 results. Therefore, it was necessary to find a method of restricting the search. During the preliminary search, using the *advanced search* function, it was noticed that a large number of trials combined a variety of therapeutic interventions. For example, warm needle acupuncture was compared to basic needle acupuncture but both groups also receive tuina massage or electro acupuncture. To prevent these kind of mixed therapeutic intervention trials being captured a number of NOT clauses were included ([box-IV.4](#)).

Table-IV.9: Search field abbreviations CNKI specialist search

Abbreviation	Chinese	Translation
SU	主题	theme, subject, topic
TI	题名	title
KY	关键词	keyword
AB	摘要	abstract
FT	全文	full text
AU	作者	author
FI	第一责任人	co-ordinator/ First investigator
AF	机构	organisation
JN	中文刊名'&'英文刊名	Chinese Journal name and English Journal name
RF	引文	citation/ reference
YE	年	Year
FU	基金	Fund
CLC	中图分类号	Chinese Library Classification
SN		ISSN
CN	统一刊号	Chinese ISSN
IB		ISBN
CF	被引频次	Citation Frequency

The *specialist search* retrieved a greater number of relevant articles. However, it was noticed that at least one article that had been retrieved using the *advanced search* had not being identified by the specialist search method. Further investigation revealed that this due to an issue connected to the structure of the Chinese language. Chinese words are commonly composed of two characters, both characters have their own meaning and can be combined with other characters to form different words. For example *jiandan* 简单 means ‘simple’; *fangbian* 方便 means ‘convenient’. The two characters *jian* 简 and *bian* 便 can be

combined to form another word *jianbian* 简便, which has the combined meaning of simple and convenient. The word for acupuncture is *zhenjiu* 针灸, which combines the characters *zhen* 针 meaning 'needle', and *jiu* 灸 meaning 'moxibustion'. Whilst Chinese words composed of one or three characters are not uncommon the majority of words are formed with two characters. Warm needle acupuncture is commonly written in a two character form *wenzhen* 温针, *wen* 温 meaning 'warm', combined with the character for 'needle' *zhen* 针. However, some authors used a 3 character form *wenzhenjiu* 温针灸. It was initially thought the two character form would be sufficient to retrieve all relevant articles, this was not the case.

A simplified search has been conducted to illustrate the problem. The search set out below aimed to retrieve all articles that contained *wenzhen* 温针 within the title as well as one of 3 terms that mean knee pain or osteoarthritis.

TI = '温针' AND TI = ('膝痹'+'膝痛'+'关节炎')

However, it did not retrieve an article by Huang Cx (2007). The title of which was:

温针灸治疗膝关节行骨性关节炎52例

If the search function treated each character as a separate lexical item, then the Huang Cx (2007) article ought to be retrieved. Note that all five characters, highlighted in red, are contained within the title. Moreover, the characters are in the correct grammatical order. However, it seems that the Chinese word for acupuncture *zhenjiu* 针灸 was treated as an individual word. The article was not retrieved due to the fact that the search did not contain the character for moxibustion, *jiu* 灸. To demonstrate the search was limited to 2007 and returned 9 results: see screenshot ([figure-IV.2](#)); the Huang Cx (2007) 黄朝曦 study was not retrieved.

The original search used three search fields - theme, subject, topic (SU) abstract (AB) and title (TI). Using this method a total of 353 results were returned when searching in the subject field. The abstract and title fields a total of 226 and 160 results were found respectively (box-IV.4). The results from the abstract and title field searches were all captured by the subject field search. Therefore only SU was used when repeating the search.

Box-IV.4: Subject, abstract and title search terms

SU=(温针灸+温针+温针刺+温针治疗) AND SU=(膝痹+膝痛+膝关节) AND
FT=(随机+对照) NOT TI=(注射+电针+玻璃+中药+推拿)=226 studies retrieved

AB=(温针灸+温针+温针刺+温针治疗) AND AB=(膝痹+膝痛+膝关节) AND
FT=(随机+对照) NOT TI=(注射+电针+玻璃+中药+推拿)=160 no additional studies

TI=(温针灸+温针+温针刺+温针治疗) AND TI=(膝痹+膝痛+膝关节) AND FT=(随机+对照) NOT TI=(注射+电针+玻璃+中药+推拿)=80 no additional studies

The research was first conducted in November 2014 and then again in February 2017. Instructions on how to use the specialist search of the CNKI database can be found at

http://epub.cnki.net/kns/help/help.aspx?helpType=zhuanye&url=help_yufa.htm.22

Figure-IV.2: Screenshot illustrating Chinese language search term issue

高级检索 专业检索 作者发文检索 科研基金检索 句子检索 文献来源检索 跨库选择

检索表达式语法

检索文献 结果中检索

发表时间：从 到

可检索字段：
 SU=主题, TI=题名, KY=关键词, AB=摘要, FT=全文, AU=作者, FI=第一责任人, AF=机构, JN=文献来源, RF=参考文献, YE=年, FU=基金, CLC=中图分类号, SN=ISSN, CN=统一刊号, IB=ISBN, CF=被引频次
 示例：
 1) TI='生态' and KY='生态文明' and (AU % '陈'+ '王') 可以检索到篇名包括“生态”并且关键词包括“生态文明”并且作者为“陈”姓和“王”姓的所有文章；
 2) SU='北京' and FT='环境保护' 可以检索到主题包括“北京”及“奥运”并且全文中包括“环境保护”的信息；
 3) SU=('经济发展'+ '可持续发展') and CF='泡沫' 可检索“经济发展”或“可持续发展”有关“转变”的信息，并且可以去掉与“泡沫”有关的部分内容。

分组浏览：来源数据库 学科 发表年度 研究层次 作者 机构 基金 免费订阅 定制检索式

2015(2) 2014(19) 2013(13) 2012(13) 2011(20) 2010(16) 2009(7) 2008(7) 2007(9) 2006(5) X
 2005(5) 2004(9) 2003(4) 2002(2) 2001(1) >>

排序：主题排序 发表时间 被引 下载 切换到摘要 每页记录数：10 20 50

(9) 清除 导出 / 参考文献 分析 / 阅读 找到 9 条结果

题名	作者	来源	发表时间	数据库	被引	下载	预览	分享
1 温针透刺治疗膝骨关节炎的临床研究	陈艺; 张包斌	实用中医杂志	2007-04-15	期刊	10	58		
2 谈蒙医温针加康复训练对膝关节骨性关节炎功能恢复的影响	于红	赤峰学院学报(自然科学版)	2007-04-25	期刊	7	82		
3 蒙医温针治疗风湿性关节炎体会	韩永华	中国民族医药杂志	2007-04-15	期刊	4	50		
4 温针治疗类风湿性关节炎	张惠	包头医学	2007-06-30	期刊	9	75		
5 温针配合穴位注射治疗风湿性关节炎146例	张均安; 杨庆林	实用中医内科杂志	2007-07-25	期刊		42		
6 温针、中药及医疗体操综合治疗原发性膝关节骨性关节炎的观察	陈国胜	当代医学(学术版)	2007-05-20	期刊	1	79		
7 温针治疗膝痛48例临床观察	黄东勉	中华中医药学刊	2007-12-10	期刊	7	58		
8 温针配合药物治疗膝骨性关节炎60例	涂成文; 李万瑶; 喻国雄	中国针灸学会临床分会第十五届全国针灸学术研讨会论文集	2007-08-01	中国会议		20		
9 温针配合药物治疗膝骨性关节炎60例	涂成文; 李万瑶; 喻国雄	世界针灸学会联合会成立20周年暨世界针灸学术大会论文摘要汇编	2007-10-01	国际会议		8		

找到 9 条结果

IV.7. Discussion

The review of the existing evidence-base has shown that acupuncture in its entirety is effective for osteoarthritis of the knee. The review also demonstrates that the physical variables of acupuncture play a role in the overall effectiveness ([section-IV.2](#)). Nevertheless, it is difficult to assess the clinical value of the calculated effect size for acupuncture versus sham controls. This is because two important issues have been not fully addressed; namely, an assessment of the external validity of the acupuncture procedures and the risk of bias from the sham protocols ([section-II.5.4](#)). Data on the acupuncture and sham procedure protocols gathered within this review inform the translational analysis ([chapter-VII](#)). This includes a novel method of evaluating external validity and risk of bias.

The systematic review identified 34 relevant clinical studies, which collectively had 3134 participants. Given the scale and number of controlled trials it would perhaps be expected that a conclusion could be drawn, with a reasonable amount of confidence. Unfortunately, this is not case. The standards of reporting are low, particularly with regards to the methods used to minimise bias. The review only considers studies published in English or Chinese. It is possible that a review without language restrictions would have retrieved more rigorous studies. Japanese and Korean styles of acupuncture use methods of moxibustion similar to TCM. Japanese and Korean databases may contain relevant trials.

Generally, the outcome measures employed do not appear to meet acceptable international standards. The citations to direct readers to further information, regarding these outcome measures, frequently lead to greater confusion rather than illumination. In conclusion, there is no reliable evidence regarding the additional use of moxibustion for osteoarthritis of the knee. There is a gap in the evidence base.

This chapter along with the narrative review collectively form phase 2 of this project. The overarching objective of phase 2 is to evaluate whether osteoarthritis of the knee is a suitable condition for the prospective RCT (section-II.5.2). All the criteria have been met (box-IV.5).

Box-IV.5: Summary of phase 2 - criteria for prospective RCT

- ▶ A semi-flexible protocol, based on the standard theory, should have a high degree of external validity
- ▶ There is evidence that acupuncture is effective for osteoarthritis of the knee
- ▶ There is evidence that the physical variables play a role in clinical outcomes for painful conditions
- ▶ There is insufficient evidence regarding the hypothesis that the additional use of moxibustion leads to greater clinical benefit

V.

Expert Interviews

采访专家

V.1. Introduction

This chapter reports on interviews conducted with experts in the field. The interviews are the first part of phase 3 of this project. Phase 3 consists of two related pieces of primary research: expert interviews and a practitioner survey. This project is guided by the framework established in [chapter-II](#), which advocates a translational research approach between four domains: TEAM, practice, clinical and mechanism research ([section-II.4.6.b](#)). Phase 3 investigates the domain of practice. Its primary objective is to map the physical components of acupuncture for osteoarthritis of the knee.

One of the main reasons for choosing expert interviews as a method is that they have knowledge of process (Flick, 2009). The researcher also has knowledge of the process of acupuncture. Acupuncture is diverse and evolving ([section-I.3](#)), so relying solely on a single practitioner's knowledge may lead to bias. The purpose of the interviews is to broaden the understanding of acupuncture for osteoarthritis of the knee, beyond that already established in the literature review ([chapter-III](#)). The second part of phase 3 is a practitioner survey. The information obtained through these interviews was used to support the development of the practitioner survey questionnaire.

Typically qualitative research is aimed at the understanding, experiences, feelings and beliefs of the participants (Creswell, 2008). However, the primary focus of this project is the physical variables of acupuncture ([section-II.2.2](#)). Consequently, the interviews sought to gather information regarding what practitioners actually do, rather than their beliefs, experiences or feelings.

The chapter discusses the development of the interview schedule, the selection of interviewees, the method of analysis and ethical considerations. In the results section new themes emerge. The conceptualisation of the clinical reasoning process developed in [chapter-II](#) is refined. The chapter considers which styles

of practice can be most readily replicated in clinical research. The limitations of the study are also discussed.

Box-V.1: Objectives of the expert interviews

- ▶ Further develop the understanding of the components of acupuncture
- ▶ Support the development a questionnaire for the practitioner survey
- ▶ Gather information on acupuncture in practice for the translational analysis of acupuncture for osteoarthritis of the knee (chapter-VII)

V.2. Methods

V.2.1. The interview schedule design

The interview schedule ([appendix-D.2](#)) was developed with reference to three sources of literature. The first was the qualitative research methodology literature that examines interviewing (Flick, 2009, Wengraf, 2001). The second was the reporting criteria for clinical trials (MacPherson et al., 2010). These criteria strive to provide a clear description of the relevant components of acupuncture, but are non-specific. The the source of literature was the narrative review presented in [chapter-III](#). which enabled the question to be tailored more specifically to the treatment osteoarthritis of the knee.

Semi-structured interviews are characterised by an emphasis on relatively open questions. Open questions help to avoid the bias stemming from the interviewee giving the answer that they feel the interviewer wishes to hear (Wengraf, 2001). The central question of the interviews was:

Please describe, in as much detail as possible, the process of treating osteoarthritis of the knee with acupuncture; from the moment a person walks into your consulting room to the end of their course of treatment.

One of the common difficulties of expert interviews is time limitation, therefore the interview should remain focussed (Flick, 2009). There was a danger that the conversation could drift from the process of treatment to the philosophical ideas

of different styles. To help retain focus on the physical processes, and ensure all relevant aspects were touched upon, a series of prompts were developed ([appendix-D.2](#)). A practice interview was carried out, as recommended (Wengraf, 2001).

V.2.2. The interviewees and sampling

Interviewees were selected by sampling representationally. This method tries to select a sample which is representative of the population (Mason, 2002). In this case, each interviewee represented a style of practice. As a lecturer in acupuncture, the researcher was familiar with the different professional organisations and teaching institutions in the UK. Four professional organisations were identified. The British Medical Acupuncture Society (BMAS); the Acupuncture Association of Chartered Physiotherapists (AACP); the British Acupuncture Council (BAcC); the Association of Traditional Chinese Medicine (ATCM). The BMAS and AACP represent those who base their practice on a biomedicine. The BAcC and ATCM represent those trained in at least one traditionally based systems of acupuncture (TBSA).

V.2.2.a. Selection criteria

The objective was that each of the main styles ([section-III.3.3](#)) and professional organisation in the UK would be represented. The eligibility criteria were as follows:

- Belong to a specific professional organisation
- Trained in and use a specific style of practice
- In current practice
- Are well-known in the field as educators or authors

Authors or educators were chosen on the assumption that their ideas influenced other practitioners. They maybe more representative of their style practice than a randomly selected individual. A total of eight individuals were interviewed ([table-V.1](#)).

V.2.2.b. *The recruitment process*

Through consultation with supervisors, and with reference to the eligibility criteria, the researcher identified suitable individuals. Initial contact was made by phone, or in person, to ascertain whether the individual would be willing in principle to be interviewed. If they agreed, then a letter was sent (via email) detailing the nature of the interview and a brief outline of the purpose of the study ([appendix-D.3](#)). The researcher travelled to a location convenient for the interviewee. Each interview was scheduled to last one hour and was recorded using an iPad. The interviews took place between November 2013 and August 2014. The longest interview was 50 minutes and the shortest 25 minutes.

Table-V.1: Interviewees - professional organisations and styles of practice

Code	Organisation	Style	Background	Practice Ys
MA	BMAS	Western Medical Acupuncture	Doctor	21
PM	AACP	Western Medical Acupuncture	Physiotherapist	12
TCM 1	BAC	TCM Contemporary		26
TCM 2	ATCM	TCM	China trained	25
PT	ATCM	TCM Five Elements	Physiotherapist	4
FE	BAC	Five Elements TCM		15
JA	ATCM	Japanese TCM		10
SB	BAC	Stems and Branches		13

The practice interview was with PT. PT's data was included in the analysis for the following reasons: PT was a physiotherapist who had trained in both TCM and Five Phase acupuncture, it was considered beneficial to use the information from someone with this background; it proved difficult to arrange an interview with a second physiotherapist AACP member; no changes were made to the

interview process; even though it was a practice interview this did not appear to affect the quality of the information.

V.3. Analysis

V.3.1. Methodological approach - saliency analysis

The methodological approach adopted to investigate the information obtained from the interviews was saliency analysis. Saliency analysis is a re-conceptualisation of thematic analysis. Saliency analysis not only assesses codes by their recurrence (themes) but also by their importance (Buetow, 2010). The advantage of this approach is that it enables comments by a small number of participants to be afforded relevance. The development of the interview schedule was based on pre-existing literature ([section-V.2.1](#)). Consequently, it was expected that the main themes were already known. A key purpose of the interviews is to discover those components of acupuncture that maybe less commonly employed. Components the researcher had either overlooked, or was unaware of as they belonged to a non-TCM style. A technique used by just one interviewee may represent an important finding, which should be included within the survey questionnaire. The aim is not only to establish the common aspects of acupuncture treatment but also to discover the range of techniques being employed. Hence, saliency analysis is ideal for the objectives of the study. The information obtained from the interviews was analysed using the framework matrix function of NVivo 10.

V.3.2. The framework method

The framework method is flexible, systematic and not aligned to any epistemology (Gale et al., 2013, Ritchie and Spencer, 1995). It identifies commonalities and differences in data, before focusing on the relationship between different parts of the data (Gale et al., 2013). In this research the aim is to map the commonalities and differences between styles of practice. It does not

seek to analyse the meaning between different parts of data, as would be typical of the last stage of the framework method (Ritchie and Spencer, 1995).

The framework method can be used to link pre-existing textual information and data derived from interviews (Ritchie and Spencer, 1995). In this case, a systematic method to link the information derived from the literature review ([chapter-III](#)) with the expert interviews. Ritchie and Spencer (1995) identify five key stages in the framework method, familiarisation, identifying a thematic framework, indexing, charting, mapping and interpretation.

Familiarisation

This is simply the process of becoming familiar with the range and diversity of the material, and gaining an overview (Ritchie and Spencer, 1995). In this study the researcher conducted and transcribed all of the interviews. This provided the initial familiarisation process.

Identifying a thematic framework

The development of a thematic framework draws on priori issues as well as on emergent issues raised by the respondents themselves (Ritchie and Spencer, 1995). A number of issues were derived from the acupuncture literature ([chapter-III](#)). These formed the basis of the interview schedule. Interrelated themes emerged and were gradually refined through a process of discussion with supervisors and returning to the original data.

Indexing

Indexing refers to the process of applying the thematic framework systematically to all the data (Ritchie and Spencer, 1995). Once the thematic framework/ index had been established the original transcripts were reviewed in order to index all of the material. The index is presented in [table-V.2](#).

Table-V.2: Framework Index

Distribution	Stimulation	Course	Constructs	Miscellaneous
Local	Needle gauge	Number of sessions	Anxiety	Cupping
Distal	Insertion depth	Frequency	Sensitivity	Other therapies
Number of points	Insertion time	Spacing out	Gradual start	External factors
Adjustment	Electro-acupuncture	Discontinuation	Patient control	Lifestyle
Location method	Deqi		Practitioner belief	Contra-indications
	Manipulation		Patient belief	
	Moxibustion		Finances and time	

Charting

Charting refers to the process of rearranging the data according to the appropriate thematic reference (Ritchie and Spencer, 1995). The information is arranged into five charts: distribution, stimulation, course, constructs, and miscellaneous ([table-V.2](#)). Charts can be arranged differently depending on whether the analysis is thematic or by case (Ritchie and Spencer, 1995). In a thematic approach the ordering of cases can be linked to specific characteristics (Ritchie and Spencer, 1995). The cases are ordered according to the homogeneity-heterogeneity of point distribution of each interviewee's style of practice. MA who uses the same acupuncture points for each patient is placed at the top of the charts. SB who individualises the treatment for each patient is placed at the bottom of the charts ([table-V.3](#)).

Mapping and interpretation

This part of the analytical process can take different forms depending on the original research question (Ritchie and Spencer, 1995). In this project, the objective is to map the range and nature of the physical variables ([section-II.2.2](#)). Many of the variables were first identified in the literature review. However, new issues emerge from the interview data that were not apparent in the

literature; these are presented predominately in the Constructs theme. Further interpretation is limited to tentative associations between styles of practice and the variables. Areas for potential future research are highlighted.

V.4. Ethics

Ethical approval for this research was obtained through London South Bank University's Ethics Committee ([appendix-D.1](#)). The following documents were used in considering the ethical issues arising from conducting this research:

- The World Medical Association Declaration of Helsinki and Social Research Association: Ethical Guidelines (2003) (World-Medical-Association, 2013).
- Social Research Association: *Ethical Guidelines* (<http://www.the-sra.org.uk/documents/pdfs/ethics03.pdf>)
- British Acupuncture Council Code of Safe Practice (British Acupuncture Council, 2016).

Any interview involves a risk of psychological intrusion. However, for this research the risk of harm was low. The interviewees were experts in the field of acupuncture. The researcher could also be described as an expert in the field, representing a particular style of practice. Different styles of practice are in effect competing philosophies. Moreover, the representative organisations compete for influence and recognition within a variety of political, social and professional domains. Conflict could arise if the interviewee felt that the researcher viewed their style of practice as inferior. Furthermore, the interviewees might have suspected the researcher had an agenda to demonstrate that his style of practice was superior. It was also possible that the interviewees may have wished to denigrate the researcher's style of practice or professional standing.

Wengraf (2001) describes how interviewees will 'frame' the nature of the interview from initial contact. The framing refers to making assumptions as to

the nature of the interview and the motives of the interviewer, conflict can be minimised by a clear description of the purpose of the interview ([appendix-D.3](#)). The purpose of the interview was to understand the processes of acupuncture rather than any beliefs that may be held. Therefore, during the interviews the researcher took care to ensure that questioning remained focussed on process rather than belief. A series of prompts were developed to help ensure the interview remained focussed ([appendix-D.2](#)).

It was also possible that a participant may reveal practice that was unsafe. Under these circumstances the researcher would have discussed the matter with his PhD supervisors. If it was agreed the practice was unsafe and put patients at serious risk, then the researcher would have contacted the participant to discuss further. If it was thought necessary, the researcher would have reported the unsafe practice to the relevant professional body. Any such action would have only taken place with the agreement of the PhD supervisors. The participant would have been informed. All eight interviews were conducted in a professional manner and no conflicts arose. Furthermore, no unsafe practice was described

All information collected during the course of the research is kept strictly confidential in line with the Data Protection Act 1998. The data will be stored on a password-protected computer at LSBU until 2020. Any information that is shared with others (eg. in reports and publications) have names and addresses removed to preserve anonymity.

V.5. Results

V.5.1. Distribution

The participants describe different typical distributions of acupuncture points. There is common agreement that local points would usually be selected. There appear to be differences in the use of distal points, either to address the knee

symptoms or as part of a whole system approach. The medical acupuncturists and TCM practitioners tend to focus on local points with some distal points. The practitioners who use another TBSA, other than or as well as TCM, have more heterogenous approaches. See [table-V.3](#) for a synopsis of interviewees' responses regarding the distribution of points.

V.5.1.a. Local acupuncture points

The use of local points is a common component of practice among the eight practitioners interviewed ([table-V.3](#)). There are minor differences between interviewees. MA only uses points local to the knee. Two interviewees mentioned that on occasions local points would not be used. For JA this is for a particular stage in the treatment:

JA: If it is very painful, if I palpate the knee and those points are painful, I may not treat those points initially. I may use more distal points, commonly in the style that I use, I may use zhongfeng Liver 4.

For SB the number of treatments without local points is relatively small:

SB: It would be quite unusual for me not to needle a local point if a patient came in with OA of the knee .. off the top of my head, 80 to 90% would have local needling, as well as other points.

Therefore the data suggests that the use of local points is a consistent component of acupuncture in practice for osteoarthritis of the knee. Moreover the use of local points is seen as a priority:

TCM 1: I think my thinking processes are 1) the knee itself.....

There is heterogeneity in the number of local points used, which range from 1 to 7 or 8.

V.5.1.b. Distal acupuncture points

There is greater heterogeneity in the use of distal points. Interviewees described three different rationales for using distal points: using the meridian to treat the pain: for anxiety; and for treating the ben-root.

MA does not use distal points. The other medical acupuncturist (PM) predominately selects local points with the addition of a distal point on the affected limb. The selection is guided by the location of the pain:

MA: I might well choose some end of meridian points, Liv 3, ST44 for example, BL60. Depending on the distribution of the pain really.

For PM the selection of distal points is normally limited to addressing the pain, with a point used for anxiety on rare occasions:

PM: It [choose a distal point for another reason] is very unlikely, but you know, if they are very anxious or if they have come in and they seem particularly stressed. But I tend not to really.

The TCM interviewees similarly use distal points as a means of controlling the pain. Although, there appears to be a widening of the rationale for choosing distal points to address the ben-root:

TCM 1: I think my thinking processes are 1) the knee itself 2) channels and is there somewhere to drag what is probably going to be stagnation to, down the channel. So I might use a distal point. And then finally what is the underlying, but I would try to make my treatment choice eloquent so if it was Spleen deficient, for example, and its medial then I might use Spleen 6.

TCM 2: Just now I mentioned local points and distal points combination, from my experience, if the patient is very painful – then we really need to use local points like A-shi points I believe it is a kind of Bi syndrome right, and sometimes related to zangfu. That is why sometimes I prefer to use sanyinjiao [Sp 6] because this relates to the 3 yin meridians, the 3 yin zang right. So I think sanyinjiao is related to zangfu.

The TCM interviewees described a hierarchy of importance, starting with local points, then distal points for pain and finally treating the ben-root.

The underlying root is not seen as a priority for this particular condition:

Researcher: Do you use distal, we talked and you said the main thing is to control the pain. What about maybe an underlying, maybe a Zangfu pattern or...?

TCM 2: No to be honest, I don't use a lot of that.....But I told you at the beginning if they present with other symptoms,, you can treat the other symptoms. But I don't think it is directly related with the knee problem.

TCM 1: I think the point I am making that I don't think it is too relevant to, I mean, if you establish fairly early on that a patient, they haven't got much other pathology then it doesn't seem relevant to be going down the line of exploring their emotions, ... This is a much more mechanical problem..... It is a straightforward treatment, not like treating depression with palpitations and insomnia and major zangfu involvement.

PT may give two sequential treatments within one session. The first treatment focuses on the pain, the subsequent treatment addresses the ben-root and is based on Five-element theory:

PT: That's where I would perhaps deviate a bit more Five Elements, I wouldn't necessarily do them all at once. If I had time I would do the local ones around the knee first and I would leave those in for 20 minutes. And then if I was wanting to tonify the underlying, I guess from my initial training I would do them afterwards for about 10 minutes.

JA on the other hand described alternating styles from one session to the next:

JA:.....because I do the Japanese and the TCM. Japanese doesn't really include the typically TCM syndrome differentiations so what I would try and do is, I may alternate them, I may do a TCM treatment then a Japanese next. It depends on what is best for the patient, in terms of improvement of quality of life, in terms of how well they are moving and how much stiffness they have.

V.5.1.c. Number of points

The number of acupuncture points used per treatment ranges from 2 to 16. There is variation not only between the interviewees, but also some practitioners vary the number: FE may use from 3 to 12 needles ([table-V.3](#)). Most interviewees typically use around 12 needles per treatment, towards the upper end of the range. As practitioners may vary the number needles by a large range, the survey included questions to capture not only the average number of needles but also the most and fewest.

V.5.1.d. Adjustment

The interviews suggest that, in practice, acupuncture will be adjusted during the course of treatment. This was in line with the theory (section-I.3.5). If the practitioner feels the first few treatments do not yield the expected results adjustments are made:

SB: *When you start with your first treatment the pulses can be quite complicated it is not necessarily clear. You start with a treatment principle or aim and then you need to augment or adjust it.*

JA: *It depends on what is best for the patient, in terms of improvement of quality of life, in terms of how well they are moving and how much stiffness they have. If I find a treatment works very well I will probably continue with that, and modify it as the patient gets better.*

PM: *Adjust it [the treatment] if they have a bad response, then continue to treat adjusting each time depending on how they feel really.*

V.5.1.e. Point location method

There are two basic methods of point location. The first relies textbook descriptions to locate the point. Alternatively painful points, locus dolendi, are located by palpation of the affected area. These are known in Chinese as A-shi points, that can be translated as *ah that's it*. Practitioners may use a combination of the two, they start with the textbook descriptions but then refine the location by means of palpation.

Interestingly, both interviewees who practice medical acupuncture stated that they used the traditional anatomical locations. One might have thought that medical acupuncturist would be more likely to conceptualise acupuncture points in terms of trigger points within the muscle body, and use palpation. It was one of the traditional practitioners, TCM 2, who strongly emphasised the use of a-shi points.

Table-V.3: Distribution synopsis

	Local points	Distal points	Number of needles	Adjustment
MA	All local	None	4 per knee	Doesn't adjust
PM	Mainly local	Meridian for pain/ rarely for anxiety	6-8 in total / usually for one knee	Yes- contralateral needling/ EA if not achieving results.
TCM 1	Priority local	Meridian for pain/ Root but not priority e.g. SP 6/ does not needle opposite knee	7 around one knee/ 3-4 General points/ 14-16 in total	
TCM 2	Priority local/ Ahshi important	Meridian for pain /always opposite knee/ Root less important typically Kidney e.g. SP 6	3 or 4 local to the knee/ and a distal point	Basic needle acupuncture first. Then EA or Moxa if necessary
PT	Always some local	Meridian for pain / Kidney points SP 6-K 3/Root depends of patient and feeling	4-5 local to knee/ 4-5 distal	
FE	Always local/ a-shi points	Depends on patient/ does not normally needle opposite knee	3-10 local to knee/2 distal points	
JA	Usually local but not always	Distal points depending on pattern/ Wide range of possibilities	4 local to the knee/ about 12 points in total	Modifies treatment / moxa can be increased if pain not reduced as expected
SB	Usually local but not always	Wide range of possibilities	2-4 needles	Adjusts treatment

V.5.2. Stimulation

Gauge, depth and insertion time

In general the interviewees use 0.25 millimetre needles, but some use either 0.16 mm or 0.35 mm (table-V.4). Needles are normally either 25 or 40 millimetres in length. MA on occasions uses 60 mm needles. The depth of insertion is dependent on the size of patient and/or the actual location of the point. Some interviewees use shallow needling (less than 5 mm) JA and SB, whereas MA inserts up to several centimetres. Some interviewees adjust the depth to obtain the desired Qi sensation. Most interviewees leave the needles inserted for approximately 20-30 minutes. Only SB might insert the needles for a short period of time. See table-V.4 for a synopsis of interviewees' responses regarding stimulation.

Deqi

The majority of interviewees try to obtain the deqi sensation. Only MA and SB do not necessarily aim to achieve deqi:

TCM 1: I have decided in my own mind that it is important to get deqi; I think we are going to get a better response I think that is what the current evidence suggests

TCM 2: Always you have to get deqi. Deqi as I told you before is not to make patients [have] a very strong feeling, but patients must have some feeling. If you have no feeling then you have to manipulate. It is very important.

MA: deqi? I don't tend to do that. I have reflected on that a lot and people talk about deqi a lot. I think deqi is relevant as a physiological marker, although it is very hard to pin down because it's so subjective in the sense that it requires a patient to perceive something then articulate it

Researcher: Do you feel for knee problems that it is important to get that so called deqi sensations?

SB: No I don't. I find that the pulse changes and symptoms change, and sometimes the patient get a very mild response and barely feels. I know some practitioner like to probe with needle, to pick up and lift and thrust to get that strong sense of..... but for me, no, you can elicit a change energetically that is very palpable without having to have strong stimulation

Even though both MA and SB do not aim to achieve the deqi sensation the physical experience of the patient would be different. This is because MA uses electro-acupuncture which means a sensation will be felt, whereas SB is content for the patient to feel very little.

Manipulation

Some practitioners mentioned the use of specific needle manipulation techniques:

SB: I tend to be thinking more about elemental points rather than channels, so it is not a case of necessarily course the Qi or changing the flow of the Qi, its more about contacting that elemental aspect to the channel. It remains anticlockwise for sedating or clockwise for tonifying

TCM 1: Anti-clockwise, putting the needle in and drawing it out. And tonification is the opposite, its like winding up a watch.

Although TCM 1 places less importance on these ideas:

I have also got in my mind thoughts of reducing and tonifying that comes from earlier practice way back. I think about that less these days and more of deqi.

Electro-acupuncture

Half of the interviewees sometimes use electro-acupuncture, TCM 1, TCM 2, PM and MA. MA uses electro-acupuncture with almost every patient. The other three use electro-acupuncture cautiously. They feel it may exacerbate the symptoms or be too painful: it is turned to when manual acupuncture has failed:

TCM 2: electro [acupuncture] is quite painful..... for my personal experience, you don't immediately use electric or moxa, you try the needle first then after a few sessions if the patient's response is not good enough, then start using different things.

PM: If I feel that, if I have given two treatments and there has been no change, then I have used the other limb and there has been no change, then my last hope is to do some electro-acupuncture. Then I might try that, I wouldn't use it as a first off really.

TCM 1: Again it is purely experience, I see flare up on electro-acupuncture. So my TCM mind is thinking we have moved too much here and there is so much

heat and stagnation the electro-acupuncture has actually made it worse. Sometimes less is more

The other four interviewees do not use electro-acupuncture. A variety of reasons were given:

JA: At the beginning I used to but I don't anymore no. I felt that people didn't like the sensation. I have actually had it on my knee, but most people I have seen didn't like it.

PT: For electro-acupuncture I haven't got a machine, we barely did it at college, we had a lecture on it..... it was always looked upon as not the traditional style...

FE: I don't use electro-acupuncture..... Its not because I particularly disapprove I am not very high tech

MD, however, feels it could be more gentle:

So it doesn't matter how sensitive the patient is, with the modern machines that control current you can start sub threshold, i.e. at levels below which nerves will respond.....And therefore, ironically in sense electro acupuncture can be a more gentle cause of stimulus for the very sensitive people who the slightest touch of the needle and they are uncomfortable. And you don't have to touch them, they can be the ones to do it.

Moxibustion

Five of the eight interviewees use moxibustion, FE, TCM 1, TCM 2, SB and JA. PT and PM are unable to use moxibustion in their clinics. PM also noted that heat packs are easier to administer. Warm needle acupuncture is used by FE, SB, JA and TCM 2. Other forms of moxibustion are also used including moxa stick and thread moxa.

Table-V.4: Stimulation synopsis

Manipulation			De Qi	Electro-acupuncture	Moxibustion
MA	No	Does not try to achieve deqi	Nearly always/ local to the knee	No - not in training	
PM		Tries to obtain	If manual does not work / local to the knee	No - not possible in clinic or convenient	
TCM 1	Clockwise for tonifying anticlockwise for sedating	Important	Yes	Yes - moxa stick local area	
TCM 2		Important	Yes - but difficult in one clinic	Yes - but difficult in one clinic/ warm needle	
PT	Even or slight reducing technique	Usually stimulates	No machine	No - not possible in clinic	
FE		Gentle DeQi	No - I'm not high-tech	Yes - important / warm needle/ moxa stick local to knee	
JA		No	No - used to but stopped	Yes - thread moxa / sometimes warm needle	
SB	Clockwise for tonifying anticlockwise for sedating	Deqi for patient not important. Sensation varies between patients More important what the practitioner feels.	No - never felt the need to explore	Yes - moxa stick and warm needle	

Table-V.4: Stimulation synopsis continued

Needles		Depth	Insertion time	Moxibustion
MA	Gauge 0.25 Length 40 - 60 mm	Deep can use 6cm needle Needs to be below subcutaneous fat	30 minutes	No - not in training
PM	25 mm and 40 mm	Superficial around patella/ deeper in the muscle - approximately 2 cm	20 minutes/ 30 if possible	No - not possible in clinic or convenient
TCM 1	Gauge 0.30 mm	1.5 cun depends on size of patient	20 minutes	Yes - moxa stick local area
TCM 2	Gauge 0.25/ 25 or 40 mm	Shallow/ can be deep in muscle	20 minutes	Yes - but difficult in one clinic/ warm needle
PT	Gauge 0.25 25 - 30 mm length	1 cun Depends on size of patient	20 minutes	No - not possible in clinic
FE	Gauge 0.25 mm	Depth required to get a sensation Can be deep	20 minutes sometimes upto 40m Depends available time	Yes - important/ warm needle/ moxa stick local to knee
JA	Gauge 0.18 - 0.2 Length 40 mm	Maximum of 5mm for Japanese, 10 mm if using TCM Depends on size of patient -	20 minutes	Yes - thread moxa/ sometimes warm needle
SB	Fine gauge 0.18 - 0.20	Guided by feeling for Qi	20 minutes Some needles only inserted for 3 minutes	Yes - moxa stick and warm needle

V.5.3. Course of treatment

Number of sessions

Interviewees typically said that an initial course would be around six weekly treatments, TCM 1 and JA suggested an initial block of around 12 treatments. See [table-V.5](#) for synopsis of interviewees' responses regarding the course of treatment.

Frequency

Treatment is generally given once a week. Occasionally twice a week, but finances and time constraints prevent this from happening on a regular basis.

Spacing treatments out

In general, the interviewees do not see the initial sessions as the end of the course of treatment. The course can be extended, with sessions being less frequent, moving to once a fortnight and then once a month. The patient is part of the decision making process.

PT: I don't mean that as a cut off and then no more, I suppose if things were working, if they are going well the absolutely and they are private and they can keep on coming then brilliant. And yes you want them to be able to manage their symptoms

Researcher: And then beyond the 12 weeks, do you go for the spreading out?

TCM 1: Yes, fortnightly then down to a month

JA: I would leave it up to the patient really, again the cost does come into it as well, depending on how people can afford it. I think if they are feeling really good I think it is good preventatively to do the acupuncture. I do have patients that come and see me who've I have been seeing for years who have arthritis and they are managing fine. They see me once every 6 weeks.

MA: the likelihood is that the pain is going to come back at some point. If say, they are completely pain free and they think this is wonderful. Our guess is the acupuncture effect will start to wear off and if you get it topped up, when it wears off, then the top up acupuncture is much more effective. That one go, top up acupuncture, is much more effective than if you really let it wear off for months and let it come back and start again. It seems that, if you say leave 6 months you have to start from scratch almost. But if you get in there just as it is wearing off say, somewhere between 6 weeks and three months, if you get in there and treat it then that probably will have a similar effect to the whole course, in a sense.

Discontinuation

It was generally felt that 3 to 6 treatments would be enough to see some response. If no improvement is observed, then treatment may be discontinued. TCM 1 summarised clearly what the typical patient with osteoarthritis of the knee would be told to expect:

TCM 1: I would say to the patient: we should see some change in four treatments. And if there was no change whatsoever in four treatments I wouldn't want to go on with acupuncture. But I would also make clear that overall treatment is going to be quite long term. Then you are talking about 12 treatments, 3 months for a chronic condition.

Table-V.5: Course of treatment - synopsis

	Sessions	Frequency	Spacing out	Discontinuation
MA	Initial course 6 to 8 treatments	Weekly / very occasionally twice a week	Should come back to top up session after six weeks to three months / can be ongoing treatment	Usually in the first 3/4 sessions
PM	6 treatments Possibly once a month if private	Once every 4 to 6 days	Yes if a patient comes privately	Usually by 4 session we can tell something is going to happen
TCM 1	12 treatment over 3 months	Once a week Twice a week if possible	Possibly spreading out to once a month	If no change after 4 treatments they stop
TCM 2	10 treatments Patient then decides if they wish to keep coming	Normally once a week Better if they can come twice a week	Yes - lets the patient decide	
PT	6 treatments Possibly spreading out if they can pay			If no change after 6 treatments need to think about referring
FE	not stated	not stated	not stated	not stated
JA	Initially 10-12 treatments	Normally once a week Twice a week if bad	Spread out to once every couple of months - 6 weeks Some come for years as a way of managing pain	Should notice improvement after 5 sessions
SB	Initially 6 weeks	No more than two weeks between sessions		If no change after 6 treatments unlikely work

V.5.4. Psychological factors

The physical process is the primary focus of these interviews, yet, psychological factors have emerged as a key theme. The psychological aspects of acupuncture may influence clinical outcomes, via components such as the therapeutic relationship (section-I.4.2.d). However, the interviews suggest that psychological factors also lead to modifications in the physical process: changes in the points selected and degree of stimulation. Four factors have emerged from the interviews: the desire to minimise the anxiety; sensitivity to needling; practitioner's beliefs; patient's beliefs. See table-V.6 for a synopsis of interviewees' responses regarding the psychological factors.

V.5.4.a. Anxiety

If they judge the patient to be anxious, interviewees may adjust the treatment. This adjustment can take the form of a reduction in the number of needles or the selection of specific acupuncture points:

JA: Sometimes when people come for a treatment for the first time they are a bit nervous, anxious. So I try not to do too many things that may increase their anxiety. The main thing on the first appointment, I try and make people as comfortable as possible. If they have never had acupuncture before make sure they are comfortable they enjoy it, it doesn't leave them feeling anxious, I want them to feel relaxed.

PM: I might do Liv 3 so that they understand what is involved. Dispel the anxiety for the next time, and off they go.

It may also mean that points thought to have a calming action are selected:

TCM 2: hegu LI4 also can be helpful to reduce the pain or help them calm down and relax. This depends on the individual patient.

V.5.4.b. Sensitivity

Closely aligned to the issue of anxiety is *sensitivity to the insertion of needles*. Interviewees assess the degree of sensitivity and potentially reduced the number of needles inserted or the strength of stimulation. They may also change the points used:

FE: I have some patients who have chronic knee problems that are so sensitive to the needling, if I can get away with using three basic points it is kind of a miracle.

The stimulation can minimised or altered:

FE: But in general often the type of patient I get, they come to acupuncture - they don't like acupuncture needling and I wouldn't say I was a very brutal needler, I am a light needling person. So I tend to reduce it as much as possible

MA: If the condition is just slightly odd if there is just a little bit more sensitivity than one might expect, or skin changes, something that is making me think of chronic regional pain – autonomic aspects which isn't common at all, but if I ever think about it – the frequencies that seem to be more effective at influencing autonomic function, and we have only got very little data on this, instead of using 80 I'd use 10 or 15 [Hz]. So 2 and 10; 2 and 15. But that rarely happens with knee pain.

If points around the affected knee are considered too sensitive, points away from the knee maybe used:

JA: If it's very painful, if I palpate the knee and those points are painful, I may not treat those points initially. I may use more distal points, commonly in the style that I use, I may use zhongfeng Liver 4.

FE: It probably goes back to this thing about being a painful treatment. If somebody is in real pain and they can't stand the needling on the affected knee then I might treat the other knee.

Early treatment - accommodate anxiety

The modifications, to account for needle sensitivity and anxiety, tend to be part of the first few treatments. Patients are allowed to get used to acupuncture, rather than deliver the 'most effective' treatment to reduce the clinical symptoms, per se.

PM: [At the first session] So quite often I give them, what I lightly call a test dose, because it is only for them to have the experience, really. And I try not choose very strong points, you know, so I might choose a point somewhere in the leg ST38 something that is not going to get them leaping off the couch. So that they have some feeling, I might do Liv 3 so that they understand what is involved. Dispel the anxiety for the next time, and off they go.

MA: I usually don't get up to good levels [of electrical stimulation] by the first, but sometimes on the 2nd they are having what I think is a decent stimulus.

V.5.4.c. Practitioner Belief

FE discussed how a practitioner's beliefs could restrict the selection of distal points and the adoption of holistic approach. Thereby, altering the physical treatment the patient receives.

Local vs distal:

FE: That you don't always have to go local, maybe; but I think it takes some courage to not go local, because it is so logical to do that.

FE: ...or you can treat the opposite arm or on the leg and stuff. Even with myself I have had to get used to that idea.....

FE: [my style] is not hard core 5 element, where you would treat the constitution and expect the knee pain would resolve as a result of improving the constitution. Which I think is theoretically sound, and I would like to try it. I am just never brave enough to work that way. For other conditions yes, but maybe not this musculo-skeletal.

It is worth noting that FE's reticence to use distal points does not stem from previous experience that was unsuccessful. FE has used contra-lateral points in patients with cancer and the results were good (needle insertion being contra-indicated in the affected limb):

FE: When I have done it, it is amazing how well it works

V.5.4.d. Patient Belief

FE discussed how the use of distal points and treating holistically was sometimes difficult for patients to accept:

FE: The problems that people have, and I have this a lot with breast cancer, if you treat on the opposite arm you have to get them used to that idea, or you can treat the opposite arm or on the leg and stuff.

FE: ..interesting to explore, what treating by the constitution would do. But I am not sure patients would necessarily buy into that. There is an expectation that if I have a problem then you need to treat the problem.

PM also noted that patients sometimes are not expecting the needles to be located away from the knee:

PM: ...need to explain the reason why, because people think it is a bit queer to expose their feet when they have come in for a bad knee.

The physical treatment may change as the patient learns more about acupuncture theory. FE highlighted a difference between those patients who present for the first time with knee pain, as opposed to those who have experience of acupuncture for other conditions. The focus on the knee pain of the former can be, for FE, more challenging:

FE: So then I would begin to treat, again if it is someone who comes in cold who has never had acupuncture before then I find that more challenging. Because I find that working with musculo-skeletal they want an instant result and I think there is a kind of educative process of getting people used to acupuncture, if they haven't had it before.

FE: But if I have existing patients and I am treating them in a Five elements manner then I will incorporate that as part of my treatment with the knee.

Other interviewees did not directly mention these issues. However, PT and JA who practice more than one style, both made statements that their choice of style changes depending on their feeling:

PT: So which brings me back to whether I am feeling slightly more Five-elementary; or if the patient is screaming out as one Element type, I will try and treat more along the channels that I feel is their constitutional Element

Table-V.6: Psychological factors synopsis

	Anxiety	Sensitivity	Gradual Start
MA	More careful if person appears anxious	Take it easy with the first needle insertion/ fewer needles if sensitive	Text easy at the beginning tries not to put them off/ encouraged to increase intensity in subsequent treatments
PM	First treatment to dispel anxiety/ rarely points inserted to reduce anxiety	Tries not to use strong points on the first treatment	Yes - gives a test treatment/ progresses to a electro acupuncture
TCM 1		Thinner needles - so not to hurt people	
TCM 2	Use LI4 to help them relax	Deep needling around patella can be painful/ uses tuina as way to reduce number of needles	electroacupuncture and moxibustion come after manual acupuncture
PT	Supine for relaxation	might use opposite knee if acute	
FE	Patients often anxious uses yintang/ ear acupuncture	Reduce number of needles / Might use opposite knee if very painful/ strength of deqi a trade-off on how much the patient can take	
JA		Can use distal points if joint very painful	
SB		Does not needle very painful joint	

Table-V.6: Psychological factors synopsis continued

	Patient control	Practitioner Belief	Patient Belief	Finances Time
MA	Controls the intensity of electroacupuncture / giving choice about frequency of treatment			
PM	Patient decides length of course of treatment - if they are paying		Needs to explain why distal points are used	
TCM 1				
TCM 2	Patient decides length of course of treatment			
PT		Treatment varies upon feeling / electro acupuncture not the traditional style		Length of treatment depends on finances patient
FE		Depends on mood	an educative process	
JA	Patient decides length of course of treatment			Length of treatment depends on finances patient
SB			Does not give advice if it might put the patient off and so have a negative effect	Finances can prevent patients coming frequently

V.5.5. Finances and time

A number of interviewees stated that the frequency and number of treatments given to any particular patient was dependent on finances, and the ability to make the time for acupuncture. These are in part objective external circumstances. However, they also contain a subjective judgement as to whether acupuncture is affordable and the necessity for more sessions. The degree of credibility, the patient's expectations and attitudes towards acupuncture may all influence the decisions regarding frequency and a number of treatments.

V.5.6. Miscellaneous

Cupping

Most interviewees stated that they do not use cupping for osteoarthritis of the knee. SB and JA might use cupping but only very rarely.

Other therapies

A range of other interventions might also be employed. These include massage, reiki, Chinese herbal medicine, heat packs and heat lamp.

External factors

A few external factors were identified that may alter the physical application of acupuncture. This mostly concerned interviewees who work in the NHS. The needle retention time may be limited or the number of treatments reduced.

Lifestyle

All interviewees consider exercise and weight management as potential recommendations for changes in lifestyle. PM and TCM 2 both noted the patient should be cautioned against doing too much, if the knee started feel better. TCM 2 and FE stated it is important to keep the knee warm.

Contra-indications

No specific contraindications were identified beyond those that apply to acupuncture in general.

Table-V.7: Miscellaneous synopsis

Cupping		Other	Lifestyle	External factors	Contraindications
MA	No	Footwear	Weight loss and exercise		Unstable clotting
PM	No	Heat pack Footwear	Exercise- be careful with the knee if they feel better	Time pressures in clinic limit length of treatment / Number of treatments also limited by GP	Unstable clotting Previous bad experience
TCM 1	No		Exercise and weight management, alcohol intake		Would not use heat if red and swollen None unless or red flags
TCM 2	No	Tuina	Exercise / be careful with the knee if they feel better / keep the knee warm		None
PT	No		Exercise is really important	Number of treatments dictated by NHS	The general contraindications
FE	-	Heat lamp / ear acupuncture / massage	keep knee warm / teaching patients to use moxa		Lymphedema Varicose veins
JA	Sometimes	Heat lamp / Chinese herbs	Diet and weight management advice	Use of moxa can be limited in the clinic	Caution with diabetes Certain medications
SB	Rarely	Glucosamine / Reiki	Diet and exercise, Qi Gong		Avoid direct needling of hot and swollen joint

V.6. Discussion

The analysis of interviews with experts in the field suggest that the components of practice are broadly in line with the theory recommended in textbooks. There are two key findings: Firstly, there is a spectrum of predictability related to style of practice that has implications for clinical research. Secondly, psychological factors may play a role moderating the physical treatment delivered.

The information gathered on the stimulation theme generally confirms the understanding of acupuncture as described in [chapter-III](#). The Japanese and Five-element acupuncturists used thinner needles with shallow insertion. Similarly, the miscellaneous theme produced no particular points of note. The literature has relatively sparse information on the course of treatment, number of sessions and frequency ([chapter-III](#)). A particular point of interest is that most interviewees stated that patients come back for top-up treatments, once every 4 to 6 weeks.

V.6.1. Replicability

The pattern differentiation process can be divided into stages of increasing complexity, as the complexity increases the potential for replicability decreases ([section-II.4.5](#)). In the narrative review of theoretical approaches, four *modes of point selection* are identified: local, meridian, general health and differential pattern points. Of these four modes, the first three can be included in a semi-flexible protocol. The differential pattern points, on the other hand, are not predictable and cannot easily be accommodated within a replicable protocol ([section-III.5.5a](#)). Consequently, one of the objectives of phase 3 is to investigate how important differential pattern points are in practice.

In general, the use of local points is the priority. Meridian points are also commonly selected ([table-V.3](#)). The medical acupuncturists concentrate on local point with only occasional meridian points. Both TCM practitioners suggested that differential pattern points were not necessarily important. They also both identified the same acupuncture point *sanyinjiao* SP6 to support the general

health. This suggests that semi-flexible protocols can be developed that would have external validity for medical acupuncture and to some extent for TCM, as well. Although, both TCM practitioners did include caveats to say, if there are other signs and symptoms these should be treated.

The practitioners who use non-TCM styles appear to include differential pattern points more commonly within their treatment strategies. Thereby, making it harder to develop a replicable protocol that would have external validity. It is not possible to predict what element PT would ascribe to any particular patient. Nor is it be possible to know whether JP would choose to treat a particular patient using Japanese acupuncture or TCM. SB always treats the patient as a whole. Therefore, apart from knowing that one needle would be placed close to the knee, it is not possible to predefine the potential acupuncture points.

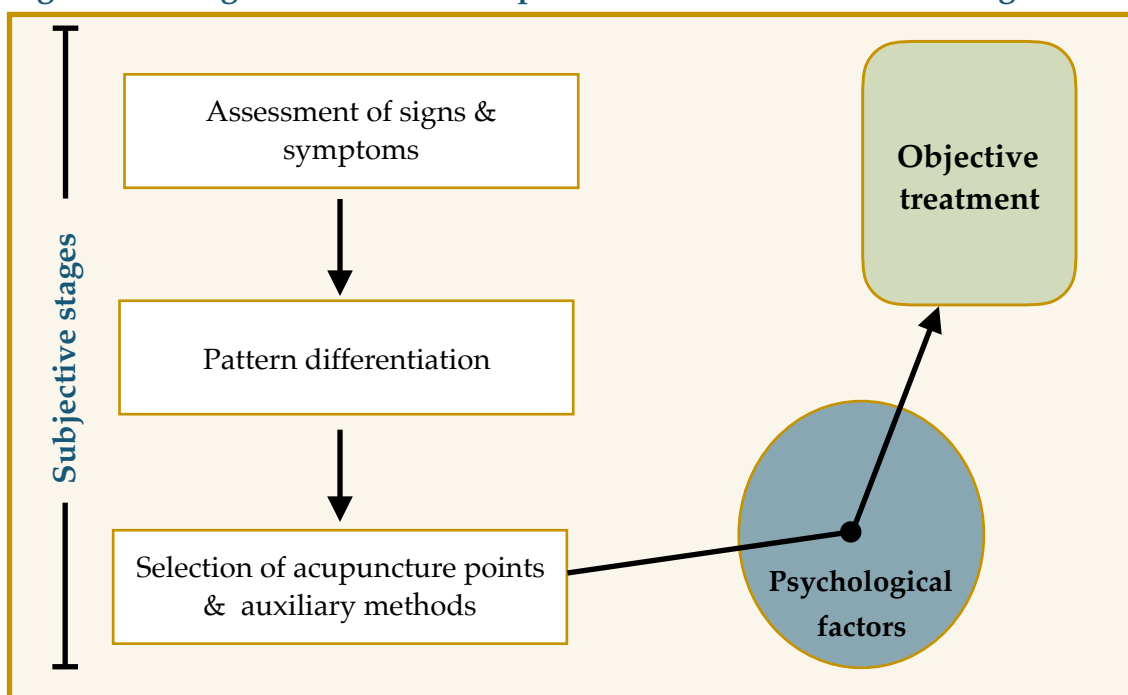
The interviewees have been placed on a spectrum of the potential to replicate their practice. MA at one end of the spectrum - high replicability, and SB at the other end of the spectrum - low replicability. The interviewees are ordered using this spectrum in the [tables-V.3/7](#). PT, FE and JP appear to vary their treatments. However, the interviews did not cover how frequently they give the more predictable TCM treatments compared to Five-element or Japanese acupuncture. Similarly, it is unclear how often the TCM practitioners address other signs and symptoms. In terms of understanding the generalisability of clinical trials, the information derived from these interviews is not adequate. The practitioner survey provides more robust information regarding that generalisability of semi-flexible protocols ([chapter-VI](#)).

V.6.2. Emergent theme - psychological factors

A group of related factors have emerged from the interviews that were not identified in the literature review ([chapter-III](#)). These form the psychological factors theme. The research framework identifies three stages of subjective decision-making prior to the implementation of physical treatment (section-II. 3.3/[figure-II.5](#)). The clinical reasoning process appears to more complicated than originally conceptualised. The physical treatment, points selected and degree of stimulation, is dependent on a highly dynamic process. The patient's real-time responses to the beginning of the physical treatment, or even the discussion of what the treatment may entail, can lead to adjustments in the physical treatment. If the patient appears anxious or sensitive after the first needle insertion there may be changes in the point prescriptions. This dynamism may be further complicated by the practitioner's perception of what the patient would consider strange or unexpected. [Box-V.2](#) summarises the additional clinical reasoning components that emerged from the interviews. The visualisation of the treatment process has been updated ([figure-V.1](#)).

Box-V.2: Additional psychological components of clinical reasoning

- ▶ Select acupuncture points to calm or relax the patient
- ▶ Modify the treatment during the first one or two sessions to allow the patient to get used to acupuncture
- ▶ Reduced the number of needles because the patient appears anxious
- ▶ Reduce the number of needles because the patient is sensitive to needling
- ▶ Adapt treatment because the patient may perceive distal points as illogical

Figure-V.1: Stages of the treatment process - enhanced understanding

V.6.3. Limitations

One of the limitations of saliency analysis is that it is dependent upon the subjective view of the researcher to identify the important issues. FE is the only interviewee who spoke at length about patients finding it difficult to accept certain forms of treatment. Similarly, it was only FE who discussed the challenges for practitioners in believing that treating the underlying condition would be appropriate for osteoarthritis of the knee. Other interviewees alluded to these issues, for example PM noted that some patients thought it was ‘a bit queer’ to insert needles in the foot for a knee problem. FE’s comments resonate with the researcher’s experience and therefore feature prominently in this analysis. It is possible that another researcher would view these comments differently. However, some of the factors were mentioned by most interviewees. In particular, that anxiety and sensitivity can lead to some adjustments in the treatment, and that this is more likely to happen in the first few sessions.

The interviews were conducted with the aim of understanding the physical components of acupuncture, in line with the ethos of the project (section-II.2.4). An approach that explored more deeply the beliefs that underpin clinical reasoning may have led to a more nuanced understanding. However, it is also

possible that the focus on the physical variables brought into sharper contrast the realisation that psychological factors partially guide the physical process.

The expert interviews fulfilled their primary objective of developing a better understanding of acupuncture for osteoarthritis of the knee. This enabled the design of the questionnaire of the practitioner survey to be enhanced. The key findings of the interviews have been summarised in [box-V.3](#).

Box-V.3 Key findings - of the expert interviews

- ▶ Components of practice appeared to be broadly in line with the theory recommended in textbooks
- ▶ The capacity to design a replicable semi-flexible protocol that has a external validity depends on the style practice. There is a spectrum with medical acupuncture having the highest replicability at one end, those who practised non-TCM styles the lowest, the TCM practitioners were in the middle
- ▶ The clinical reasoning process includes psychological factors - these had not been identified in the narrative review ([chapter-III](#))

VI.

Practitioner Survey

医生调查

VI.1. Introduction

This chapter reports the second piece of primary research within phase 3 of this project, a survey of acupuncture practitioners. The key objective of the survey is to gather information on the practice of acupuncture for osteoarthritis of the knee; in particular the physical process ([section-II.5.3](#)). As far as the author is aware, this is the first practitioner survey to be carried out specifically for osteoarthritis of the knee. The reference lists of clinical trials identified in the systematic review were studied ([chapter-IV](#)) and specific searches undertaken using: Medline, CINAHL complete, AMED, PsycINFO and Scopus ([appendix-E. 20](#)); yet, there appears no similar previous research.

This chapter discusses the development of the survey questionnaire. The structure of the questionnaire evolved from the charts established in the preceding chapter ([V](#)). There are five core themes: distribution (of needles), stimulation methods, course of treatment, adaption of treatment and miscellaneous methods. The questionnaire also covers background data on respondents along with information on safety and adverse events. The methods of distributing the survey (electronic and paper-based) as well as sampling (randomised and non-randomised) are described. The survey generated quantitative data that is analysed using the Kruskal-Wallis and Mann-Whitney tests.

The diversity of acupuncture styles and the synthetic relationship between them pose challenges for research ([section-I.3.3](#)). Therefore, the survey data is used to investigate the similarities and differences between styles of practice, commonly used in the UK. The final section of the chapter debates the key findings and limitations of the study. Knowledge developed through this survey is incorporated into the translational analysis presented in the following chapter ([VII](#)).

Box-VI.1: Objectives of the practitioner survey

- ▶ To gather data on the physical components of acupuncture and associated techniques used to treat osteoarthritis of the knee in clinical practice, in the UK.
- ▶ To evaluate the intra and inter-style homogeneity and heterogeneity of acupuncture for osteoarthritis of knee

VI.2. Methods

VI.2.1. Development of the questionnaire

The content of survey questions can be divided into five distinct types: behaviour, beliefs, knowledge, attitudes and attributes (De Vaus, 2002). The project's focus is on the physical processes of acupuncture ([section-II.2.2](#)). Consequently, the survey sought to collect information on the 'behaviour' of acupuncturists - to know what they actually do. The questionnaire consists of seven sections. The 'behavioural' information was collected in the five core sections of the questionnaire (A,B,C,D,E). The first section (Background) collected 'attribute' information on the practitioners themselves. The final section contains two open-ended qualitative questions. The first, allowed respondents to describe any additional components of practice; to capture any techniques that had not been covered in sections A-E. The second open-ended question requests information on adverse events.

The majority of questions in the survey are multiple choice and use a grid format for simplicity and ease of use (De Vaus, 2002). Each grid uses one of two basic stems: "The proportion of chronic knee pain treatment sessions for which:" and "The proportion of chronic knee pain patients for which:" These are followed by specific statements such as "I insert needles at local points" or "I use a moxa stick". The answer options are ordinal categories (De Vaus, 2002): never, less than 25%, 25 to 75%, more than 75%, always ([appendix-E.7](#)). These

were preferred to adverbs such as ‘often’ or ‘seldom’ so respondents would be encouraged to think in numerically meaningful terms. Other questions were either multiple-choice or required a simple numerical answer.

There were 61 questions, in total

- Background information (6)
 - A. Point distribution (15)
 - B. Stimulation (18)
 - C. Course of treatment (5)
 - D. Adaption of treatment (6)
 - E. Miscellaneous (9)
- Safety and further information (2 free text questions)

The nosology of TBSA and biomedicine are different ([section-I.3.5.a](#)). Therefore to allow practitioners to consider the symptomatic picture as they would in practice the following brief clinical scenario was provided:

- The typical case is someone over 50
- They suffer from chronic pain and stiffness in one knee.
- They have been told by their GP they have osteoarthritis of the knee.
- The knee does not feel hot to touch or to the patient. The pain is not made worse by warmth.
- There has been no recent fall or twisting of the joint.

The design and contents of the questionnaire are based on: the research framework ([chapter-II](#)); literature review of acupuncture for knee pain ([chapter-III](#)); the expert interviews ([chapter-V](#)); the recommendations for reporting acupuncture in clinical trials (MacPherson et al., 2010).

To ensure content validity it is recommended that surveys should be given to a panel of experts who can make suggestions for the adequacy and relevance of the questions (Bowling, 2009, Parahoo, 1997). Draft versions of the

questionnaire were submitted to members of the Confucius Institute Research Collaboration (CIRC) (n=9). CIRC are all acupuncturists and many have extensive knowledge of research methods. The researcher is primarily trained in the TCM style of practice, this posed a risk for the development of the questionnaire. The potential bias was ameliorated by: the literature review ([chapter-III](#)) encompassing non-TCM texts; the interviews included non-TCM experts ([chapter-V](#)); some members of CIRC practised non-TCM styles. A qualitative question was included to capture any methods that had not been identified through the development process. The feedback from CIRC confirmed that the basic structure of the questionnaire was adequate and no major alterations were required.

Section A is concerned with how the needles are distributed over the body, it contains two parts. The first part asks questions regarding the number of needles used (Q7-12). The grid question (Q13) uses the stem “The proportion of chronic knee pain treatment sessions for which:” It consists of nine questions based on the *modes of point selection* identified in [chapter-III](#) and [chapter-V](#) ([table-VI.5](#)). Two of the questions relate to selecting points to benefit the general health of the patient, and not simply the pain. General health (Q13f) refers to the use of well-known acupuncture points that are thought to benefit general health. Other symptoms (Q13h) is potentially more complex and would probably be accompanied by the use of a TBSA theoretical model. The fundamental objective is to understand whether acupuncturists are using points that could not be predicted from the signs and symptoms associated with clinical osteoarthritis. If the questionnaire had been directed purely at TBSA practitioners, it could have referred to selecting points to address the ben-root differential pattern ([section-III.4.3](#)). This may not have been meaningful for some medical acupuncturists. Therefore, it was decided to frame question simply in terms of ‘other signs and symptoms’. With the intention that this would be meaningful to all practitioners and did not inadvertently imply a belief in a TBSA.

Section B looks at ways in which the acupuncture points are stimulated. There are three questions regarding the basic parameters of needling: depth of insertion, needle gauge, and retention time. There are also questions on the use of manual needle manipulation, electro-acupuncture, intra-dermal needles and moxibustion ([table-VI.9/10](#)). The questions on moxibustion and electro-acupuncture follow the modes acupuncture point selection ([section-III.5.5.a](#)) in differentiating local, meridian and wider body acupuncture points.

Section D enquires about ways in which treatment maybe adapted. There were four questions on psychological factors that could influence the treatment, these were identified by the expert interviews ([chapter-V](#)). Two further questions are asked regarding the practical considerations of time and finance.

Section C consists of five questions on the number and frequency of treatment sessions. Section E has six questions on the use of auxiliary techniques: cupping, heat lamp, massage and TENS; as well as three questions on the advice practitioners may give.

VI.2.2. Sampling frame

The sampling frame was identified as the 12000 practitioners who belong to one of the four professional organisations outlined in [table-VI.1](#). Ideally a sampling frame should contain a complete list of every element in the target population (Bowling, 2009). It is possible that some practitioners belong to more than one register, which would slightly increased their chance of being selected. Some practitioners may chose not to be listed on the public database. There are a number of practitioners who do not belong to professional organisations. However, this research is concerned with acupuncture as practised by those who have a professional qualification. It is thought relatively few of those with professional qualifications would choose not to belong to a professional organisation. None of these scenarios were considered to pose a significant risk of bias

Table-VI.1: Sampling frame - Professional Organisations

Organisation	Membership*
British Acupuncture Council (BAcC)	3000
British Medical Acupuncture Society (BMAS)	2000
Association of Traditional Chinese Medicine (ATCM)	800
Acupuncture Association of Chartered Physiotherapists (AACP)	6000

* data taken from the organisations' websites

VI.2.3. Distribution of the survey

Data collection was between December 2015 and April 2016. Four different methods were used to distribute the survey: a non-randomised promotion of a Survey-monkey link; a snowballing approach via contacts of the researcher; letter and e-mail sent to random samples of practitioners. Financial restrictions ruled out the possibility of providing any incentive to complete the survey.

Method 1: Promotion via professional organisations

The survey was distributed by electronic means using Survey-monkey. A link to the survey was included in the electronic newsletters, and/or posted on the members website, of the four organisations identified in [table-VI.1](#).

Method 2: Snowballing emails to contacts

Emails were sent to acupuncture practitioners known to the researcher requesting them to complete the survey. The emails also contained a request to forward the link to other practitioners ([appendix-E.3](#)).

Methods 3: Letter format - randomised

One hundred practitioners, 25 from each organisation, were randomly selected from the publicly available registers. The questionnaire was mailed to them along with a stamped addressed envelope, introductory letter ([appendix-E.4](#)), participant information sheet ([appendix-E2](#)).

Method 4: Email format - randomised

One hundred practitioners, 25 from each organisation ([table-VI.1](#)), were randomly selected from the publicly available registers. A personal email was sent by the researcher requesting that they complete the survey by using link to

Survey-monkey embedded in the email ([appendix-E.4](#)). The participant information sheet was included in the email ([appendix-E2](#)).

VI.2.4. Randomisation

Randomisation was carried out in the week beginning 16th November 2015. The names of all practitioners on the BAcC, ATCM and BMAS websites were downloaded and put into spreadsheets. Due to differences in the way the contact information was arranged it was necessary to manually remove duplicates from the ATCM and BMAS lists. For each professional body two sets of 25 random numbers were generated: one for the practitioners who would receive a postal questionnaire and one for those who would receive a personalised email. The BAcC list contained 2878 names, BMAS 260, and ATCM 667. Random numbers were generated using an online generator (Urbaniak and Plous, 2013). Practitioners were selected according to their position on the list; for example, if 127 was generated then the 127th name on the list was selected.

The website for the AACP randomly selects practitioners based on certain criteria (AACP, 2014). This method was used selecting a single criterion, those who treat 'musculoskeletal problems'. This was preferred to downloading the entire register, as it was more convenient and increased the likelihood that the practitioner would be treating patients with osteoarthritis of knee.

For practitioners chosen to receive an email: if an email address was not present on the register, a brief online search was carried out for the practitioner's professional website. If the website could not be found; or there was any doubt that the practitioner was the individual selected through randomisation; or the website did not supply an email address, then the practitioner was discounted. The randomisation process was then repeated to select another practitioner. The emails were sent in the week beginning 7th of December to the practitioners who had been randomly selected. If an email bounced, a new name was selected using the same randomisation method. This was a particular problem for practitioners on the AACP register. This maybe have been due to the fact

that many of them work in organisations such as the NHS where external emails could be treated as spam (Bishop et al., 2011c).

For practitioners chosen to receive a letter: if a full postal address was not provided on the register, a brief internet search was conducted to find the practitioners professional website. Only if the website clearly belonged to the individual randomised, then the postal address was taken from the website to complete the missing information. When this was not possible the practitioner was discounted and the randomisation process repeated. The paper copies were posted between the 3rd-8th of December. One questionnaire was returned unopened to the researcher the following week. A new name was randomly selected from the register and the questionnaire resent.

VI.2.5. Anticipated responses

It was anticipated that Methods 1 and 2 would have a very low response rate. An estimate of 2% of the sampling frame would lead to 240 completed questionnaires. Neither of these methods involved randomisation and were consequently at risk of bias. The list of contacts of the researcher primarily consisted of traditional practitioners, and included all heads of acupuncture courses that have formal relationship with the British acupuncture accreditation board (BAAB). BAAB is a sister organisation of the BAAC. As such, the snowballing method would probably garner more responses from traditional acupuncturist than medical acupuncturists, and potentially bias the results. It was perceived that greater diversity of practice styles existed within the BAAC. So the potential bias was not considered problematic, indeed may even help in obtaining sufficient responses from practitioners of less commonly used styles, such as Japanese acupuncture. Promoting the survey by the professional registers potentially biases the response rate towards those interested in research or likely to use electronic media. To ameliorate these effects the randomised sampling methods were also used. It was anticipated that methods 3 and 4 would have a higher response rate. An estimate of 40% response rate

would lead to 80 completed questionnaires. Therefore, it was hoped that a total of 320 practitioners would complete the survey.

VI.2.6. Analysis

The data was analysed using IBM® SPSS® Statistics Version 21. Frequencies, averages and ranges that describe the practice of acupuncture were generated. The Kruskal-Wallis test was used to establish whether there were statistically significant differences in the responses between groups. If differences were detected, then follow-up pairwise comparisons using the Mann-Whitney test were conducted. A p-value of 0.05 was taken as the criterion for statistical significance.

The median scores were used as a simple method of quantifying the relative importance of various components in clinical practice. The ordinal category responses (never, less than 25%, 25 to 75%, more than 75%, always) were labelled numerically 0-4: where never = zero, less than 25% = 1, and so on, to always = 4. Thus, a median of 3 indicates that at least 50% of practitioners use that criterion more than 75% of the time. A median of 1 indicates that 50% of acupuncturist use the component less than 25% of the time.

VI.3. Ethics

Ethical approval to conduct this research was obtained through London South Bank University's Ethics Committee ([appendix-E1](#)). The World Medical Association Declaration of Helsinki and Social Research Association: Ethical Guidelines (World-Medical-Association, 2013); and London South Bank University Research Ethics Committee Code of Practice for Research Involving Human Participants (LSBU, 2011) were used for guidance. The LSBU code of practice stipulates that participants should have autonomy and have 'the right to think independently and act freely to decide to participate' (LSBU, 2011). The participants are professional acupuncturists and were considered capable of

making independent decisions as to whether to complete the questionnaire. Consent was seen as implicit if the questionnaires were completed and returned.

VI.4. Results

The results are presented in broadly the same sequence as the questions in the survey: background information; distribution; stimulation; moxibustion; course of treatment; adaption of treatment; miscellaneous; advice; safety.

VI.4.1. Response rates

There were a total of 230 responses, of these 39 were from method 3 - letter format. Out of the 230 respondents, 19 only completed the background information questions, these participants were removed from the data analysis, leaving 211 responses.

Practitioners were asked to describe their style of practice, they could select more than one option. The choices were, Five element, Japanese, TCM, Medical acupuncture and other. Those who selected other were asked to specify their style. TCM was chosen by 118 respondents, followed by Medical acupuncture, 76; Five-element, 26; Japanese, 10. There were also 17 respondents who stated they used another style. Some practitioners reported that they practiced more than one style, this was expected given the known diversity of education ([section-I.3.3](#)) Of the 26 Five-element practitioners, 18 also utilised TCM; of the 10 Japanese style practitioners, 8 used TCM. Very few Five-element or Japanese practitioners used these styles exclusively, 6 and 1, respectively. As neither the Five-element or Japanese had sufficient respondents to make a statistically meaningful group it was decided to combine the responses of these practitioners. The respondents were categorised into one of three groups, Medical, TCM and Combined based on the following criteria:

- Respondents that indicated they only used Medical acupuncture were classed as Medical.
- Respondents that indicated they only used TCM were classed as TCM
- Respondents who indicated that they used Medical acupuncture and TCM were classed as TCM.
- Respondents who indicated that they used Japanese/Five-elements/ other non-TCM TBSA were classed as Combined.

Twenty eight respondents gave qualitative descriptions of their practice style. These were manually allocated into one of the three groups using these descriptions as well as the criteria above ([appendix-E.5](#)). [Table-VI.2](#) shows the final number of respondents in each group, and the number that were manually allocated.

Table-VI.2: Number of respondents for each allocated group

Group	n	Manually Allocated
Medical	73	5
TCM	96	9
Combined	42	14

The greatest number of responses was achieved from the BAAC, 85; followed by AACP, 64; BMAS, 30; ATCM, 25. There were also 11 responses from practitioners who belonged to other professional organisations and five respondents did not complete this particular question ([appendix-E6](#)). A number of the other professional organisations were non-UK based, USA, Australia, Norway, Ireland and Portugal. It is not clear if these individuals were practising in the U.K. The original objective was to describe acupuncture in the U.K.. As there is no obvious reason why acupuncture should be different in these countries compared to the UK, these responses were included in the analysis.

VI.4.2. Background information

Respondents ranged from being in their first year of practice to 40 years of clinical experience, the median number of years in practice was 12. The TCM and Combined styles both had medians of 15 years in practice, the medical acupuncture group 10 years.

The median number of people with osteoarthritis of the knee treated each month was four, lower and upper quartiles were two and six. The Combined style tended to see fewer cases per month with a median of 2.5; compared to 4 and 4.5 for medical and TCM, respectively. Whilst there are some slight differences, the three groups appear to be broadly similar in terms of clinical experience.

Over 50% of respondents belong to statutory regulated (non-acupuncture) health care professions; the largest profession was physiotherapist ([table-VI.3](#)). As expected the BAAC and ATCM predominantly practiced TBSA. BAAC members practiced the widest range of styles, whilst ATCM respondents tended to focus on TCM. Despite being nominally a medical acupuncture association, the AACP had a significant minority of respondents who used TBSA, mostly TCM ([table-VI.4](#))

Table-VI.3: Acupuncture association by regulating bodies and style

Association →	AACP	ATCM	BMAS	BACc	Other	Total
Regulating body						
Chiropractor	0	0	2	0	0	2
Doctor	0	4	12	2	2	20
Nurse	0	1	5	4	0	10
Osteopath	0	1	4	2	0	7
Physiotherapist	66	1	6	7	2	82
Non-statutorily	0	20	0	64	7	91
Other statutorily	0	2	1	3	3	9
Style of practice						
Five-Element	3	0	24	0	0	27
Japanese	0	2	7	0	0	9
Medical	45	0	3	29	1	78
TCM	15	25	70	2	7	119
Other	6	0	8	1	2	17

AACP- Acupuncture association of chartered physiotherapists; ATCM - Association of traditional Chinese medicine; BMAS - British medical acupuncture society; BACc -British acupuncture association

Table-VI.4: Allocated style of practice by statutory regulating bodies

Profession	Medical (73)	TCM (96)	Combined (42)
Chiropractor	2	0	0
Doctor	12	3	4
Nurse	5	4	1
Osteopath	5	2	0
Physiotherapist	48	18	6
Total	72	27	11

VI.4.3. Point distribution (section A)

Section A of the questionnaire ([appendix-E7](#)) is concerned with how the needles are distributed over the body, it contains two parts. The first part asks questions regarding the number of needles used ([Q7-12](#)). The second part ([Q13](#)) uses the stem “The proportion of chronic knee pain treatment sessions for which:” It consists of nine questions ([table-VI.5](#)).

Table-VI.5: Stem Questions - Point distribution Q13

Nō	Mode of point selection	Abbreviation
13a	I use A-shi points/ reactive points	A-shi
13b	I insert needles at local points (within 10 cm of the knee)	Local
13c	I insert needles distal to the knee, on the affected leg, selected according to the meridian affected/ distribution of pain	Meridian
13d	I insert needles around the unaffected knee	Unaffected knee
13e	I inserting needles at the elbow to reduce knee pain	Elbow
13f	I use distal point to help support/ strengthen general health, such as SP6 <i>sanyinjiao</i> or K3 <i>taixi</i>	General health
13g	I use distal points to help calm/ relax that patient, such as LI4 <i>hegu</i> or <i>yintang</i>	Relaxation
13h	I insert needles at points to address other signs and symptoms (such as insomnia, indigestion, breathing difficulties etc)	Other symptoms
13i	I use Back Shu points e.g. BL11, BL20, BL23	Back-shu

VI.4.3.a. *Number of needles - Questions 7-12*

The Kruskal-Wallis test identified significant differences between groups for all six questions regarding the number of needles used except for the average number around the knee (Q12), for which the difference was not statistically significant Kruskal-Wallis $\chi^2(2) = 4.14$, $p = 0.13$ (table-VI.6). For Q12, the average number of needles around the knee, the inter-quartile ranges for both TCM and Medical acupuncturists was four to six whereas for the Combined group was three to five (table-VI.6). However, for the total number of needles in a typical treatment (Q9) TCM acupuncturists tend to use more; the inter-quartile range was seven to 11.5. The inter-quartile range for Medical and Combined groups was five to eight (table-VI.6). The medians for Medical, Combined and TCM groups were 6, 7 and 8, respectively. The Mann-Whitney tests indicated that the difference was statistically significant when TCM was compared to both Medical and Combined groups (table-VI.7).

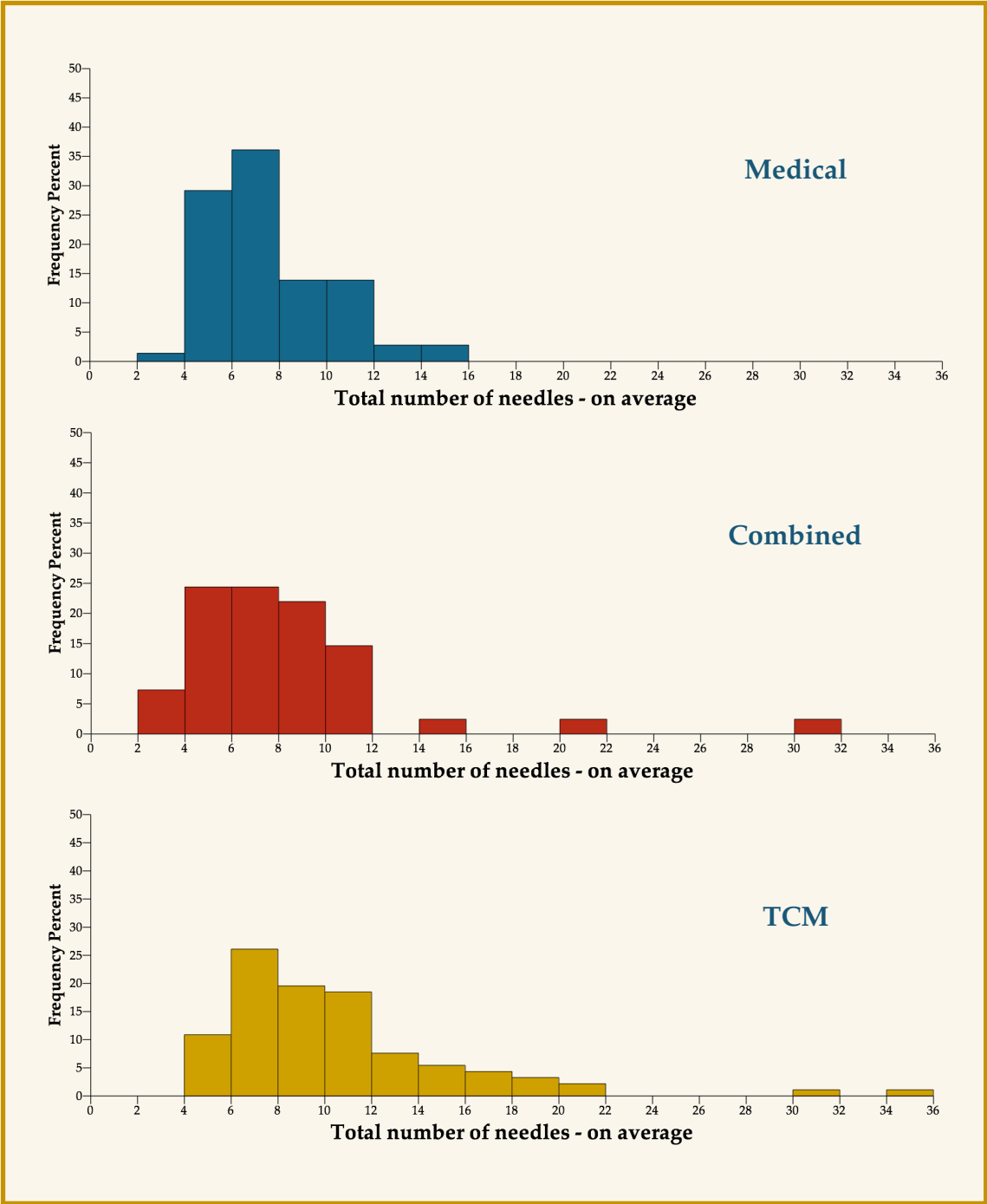
The statistical differences between groups appear to be due to different distributions within groups rather than different central tendencies between groups. The medians were often similar or even identical. For example, the median for the most needles that would be used around the knee (Q10) was six for both Medical and TCM groups, yet the Mann-Whitney test indicated a statistically significant difference between the groups $U = 2387$, $p = 0.017$. Similarly for both TCM and Combined the median for the fewest needles that would be used around the knee (Q11) was two, but the Mann-Whitney test indicated a statistically significant difference between the groups $U = 1040$, $p = 0.001$ (table-VI.7). Figure-VI.1 shows the distributions for the total number of needles used - on average. The distribution graphs of the responses to the questions regarding the most and the fewest needles used, as well as the number inserted locally to the knee are presented in appendix-E.9.

Table-VI.6: Number of needles used (Q7-12) Ranges and Kruskal-Wallis tests

Question		Group	Median & Inter-quartiles					Range			Kruskal-Wallis		
			25	median	75	IQR	Min	Max	Range	Mean Rank	χ^2	p	
7	Most needles	Medical	7	8	12	5	4	25	21	85	19.52	<0.001	
		TCM	9	12	15	6	4	70	66	123			
		Combined	6	8	12.5	6.5	3	30	27	90			
8	Fewest needles	Medical	3	4	5	2	1	9	8	93	7.77	0.021	
		TCM	3	4.5	6	3	0	20	20	116			
		Combined	2	4	5.5	3.5	1	15	14	93			
9	Average number of needles	Medical	5	6	8	3	3	15	12	83	22.67	<0.001	
		TCM	7	8	11.5	4.5	4	35	31	125			
		Combined	5	7	8	3	3	30	27	89			
10	Most local to knee	Medical	4	6	7	3	2	10	8	92	13.00	0.002	
		TCM	5	6	8	3	1	20	19	113			
		Combined	3	5	7	4	0	10	10	75			
11	Fewest local to knee	Medical	2	4	4	2	0	7	7	112	16.38	<0.001	
		TCM	2	2	4	2	0	7	7	97			
		Combined	1	2	3.25	2.25	0	6	6	68			
12	Average local to knee	Medical	4	5	6	2	2	10	8	102	4.14	0.126	
		TCM	4	5	6	2	1	10	9	101			
		Combined	3	4	5.5	2.5	0	10	10	81			

IQR - inter-quartile range $p < 0.05$ considered statistically significant- Non significant results shaded in green

Figure-VI.1: Total number of needles used - on average (Q9)



In general the TCM group had greater internal heterogeneity. For example, the inter-quartile range for the average number of needles used in total was three for Medical and Combined groups but 4.5 for TCM; the ranges were 12, 27 and 31 for Medical, Combined and TCM, respectively ([table-VI.6](#)). In comparison the inter-group differences in medians were only of the order of 1-2 needles ([table-VI.6](#)). Internal heterogeneity was not confined to the TCM group, however. A medical acupuncturist may use a maximum of only four needles, whereas a another may never use less than eight ([table-VI.6](#)). The number of needles used appears to depend on the individual practitioner rather than being a hallmark of their style of practice.

Table-VI:7 Number of needles used (Q7-12) Mann-Whitney U tests

	Medical vs TCM		Medical vs Combined		TCM vs Combined	
	U	<i>p</i>	U	<i>p</i>	U	<i>p</i>
Most needles	2055	< 0.001	1433	0.796	1298	0.004
Fewest needles	2553	0.011	1462	0.932	1486	0.049
On average	1960	< 0.001	1407	0.677	1257	0.002
Most needles (knee)	2387	0.017	1077	0.079	1040	0.001
Fewest needles (knee)	2488	0.065	751	< 0.001	1105	0.005

Asymp. Sig. (2-tailed) Non significant differences shaded $p < 0.05$

VI.4.3.b. Modes of point selection - Question 13(a-i)

The data showed that some of the modes of point selection are more frequently employed than others. To support the analysis of the results the sub-questions have been re-ordered according to the median response within the TCM group ([section-VI.2.6](#)). [Table-VI.8](#) shows the different modes with the most frequently used at the top. The Kruskal-Wallis test indicated that there was a statistically significant difference between groups for all modes except Meridian points (13.c). The follow-up Mann-Whitney tests showed that these differences were predominantly between medical acupuncturists and both TBSA groups ([table-VI.9](#)).

Table-VI.8: Q13 Stem questions - Kruskal-Wallis and medians

	Question	Allocated Style	Median	N	Mean Rank	χ^2	p
13b	Local	Medical	4	70	105	16.4	< 0.001
		TCM	4	87	101		
		Combined	3	34	65		
		Total		191			
13c	Meridian	Medical	3	69	86	5.1	0.078
		TCM	3	87	105		
		Combined	3	35	92		
		Total		191			
13f	General health	Medical	1	70	62	43.7	< 0.001
		TCM	3	86	116		
		Combined	3	34	114		
		Total		190			
13a	A-shi	Medical	2	68	74	18.0	< 0.001
		TCM	3	85	111		
		Combined	2	36	95		
		Total		189			
13g	Relaxation	Medical	2	69	80.06	8.9	0.012
		TCM	2	88	105.39		
		Combined	2	32	98.64		
		Total		189			
13h	Other symptoms	Medical	0	70	51.29	77.4	< 0.001
		TCM	2	88	119.36		
		Combined	3	34	130.40		
		Total		192			
13d	Unaffected knee	Medical	1	70	72	18.7	< 0.001
		TCM	1	84	99		
		Combined	1	27	114		
		Total		181			
13i	Back shu	Medical	0	70	66.43	33.6	< 0.001
		TCM	1	86	107.78		
		Combined	1	31	118.02		
		Total		187			
13e	Elbow	Medical	0	70	64	43.7	< 0.001
		TCM	1	82	104		
		Combined	1	31	124		
		Total		183			

df = 2

Significance $p < 0.05$ Non significant results shaded in green

Table-VI.9: Question 13 Modes of point selection Mann-Whitney tests

Question		TCM vs Medical					Combined vs Medical					TCM vs Combined				
Nō	Abbrev	N	U	p	z	r	N	U	p	z	r	N	U	p	z	r
13b	Local	157	2898	0.554	-0.592	-0.05	104	701	< 0.001	-3.696	-0.36	121	905	< 0.001	3.580	0.33
13c	Meridian	156					104					122				
13f	General health	156	1306	< 0.001	6.208	0.50	101	536	< 0.001	4.639	0.46	120	1433	0.858	0.179	0.02
13a	A-shi	153	1771	< 0.001	4.210	0.34	104	947	0.051*	1.95	0.19	121	1262	0.117	1.566	0.14
13g	Relaxation	157	2216	0.003	2.976	0.24	104	893	0.115	1.577	0.15	120	1314	0.564	0.576	0.05
13h	Other symptoms	158	793	< 0.001	8.220	0.65	104	313	< 0.001	6.416	0.63	122	1221	0.106	-1.617	-0.15
13d	Unaffected knee	154	2062	0.001	3.384	0.27	104	493	< 0.001	3.889	0.38	111	963	0.211	-1.250	-0.12
13e	Elbow	152	1584	< 0.001	5.486	0.44	97	386	< 0.001	6.151	0.62	113	987	0.054	-1.924	-0.18
13i	Back-shu	156	1685	< 0.001	5.050	0.40	101	481	< 0.001	4.857	0.48	117	1193	0.359	-0.918	-0.08

Significant differences (2-tailed) $p < 0.05$ Non statistically significant differences shaded in green r = effect size * close to statistical significance

Local points and a-shi points (Q13a-b)

Local points are an important part of practice across all three styles ([figure-VI.2:13b](#)). Only four respondents (< 2%) stated that they would never use local points. A high proportion of medical and TCM practitioners select points local to the knee for more than 75% of treatments; 85% and 90% respectively. The proportion drops to 56% for Combined styles. The use of local points is the only mode where there was a statistically significant difference between the two TBSA groups: $U=905$, $p < 0.001$. It is also the only mode that is used more by medical acupuncturists than those in the Combined group: $U=701$, $p < 0.001$ ([table-VI.9](#)).

The use of a-shi points refers to palpating around the afflicted area (local) to locate specific points of pain ([section-V.5.1.e](#)). As can be seen in ([figure-VI.2:13a](#)), the majority of practitioners in all three styles use this method of point location on at least some occasions. TBSA practitioners tend to use a-shi more frequently than medical acupuncturists, the Mann-Whitney was statistically significant when the Medical group was compared to TCM $U=1771$, $p < 0.001$; and approached significance when compared to the Combined group, $U=947$, $p = 0.051$ ([table-VI.9](#)).

Meridian points (Q13c)

The selection of meridian points based on the distribution of pain is the only mode in which no statistically significant differences were found between any of the groups ([table-VI.8](#)). These points are commonly used in practice, apart from 10% of medical acupuncturists, and one TBSA practitioner, all practitioners use these points at least some of the time. The majority do so in more than 75% of treatment sessions. However some practitioners, in each group, use these points less than 25% of the time ([figure-VI.2: Q13c](#)). There is effectively no inter-group differences, but heterogeneity within each group.

Relax (Q13g)

The majority of respondents stated that on occasions they use points to help patients relax. How frequently they did so was fairly evenly distributed across the options provided. There was a statistically significant difference between the TCM and Medical groups ([table-VI.9](#)), however, this appears to have been due to the extremes of 'never' and 'always'. Only 3% of TCM respondents would never use points for relaxation, compared to 19% for Medical. Whereas 24% of TCM always use these points, compared to 7% for the medical group. Otherwise, there appears to be little difference between the groups ([figure-VI.2: Q13g](#)).

Figure-VI.2: Distribution graphs: Q13 (a,b,c,g) A-shi, Local, Meridian and Relaxation

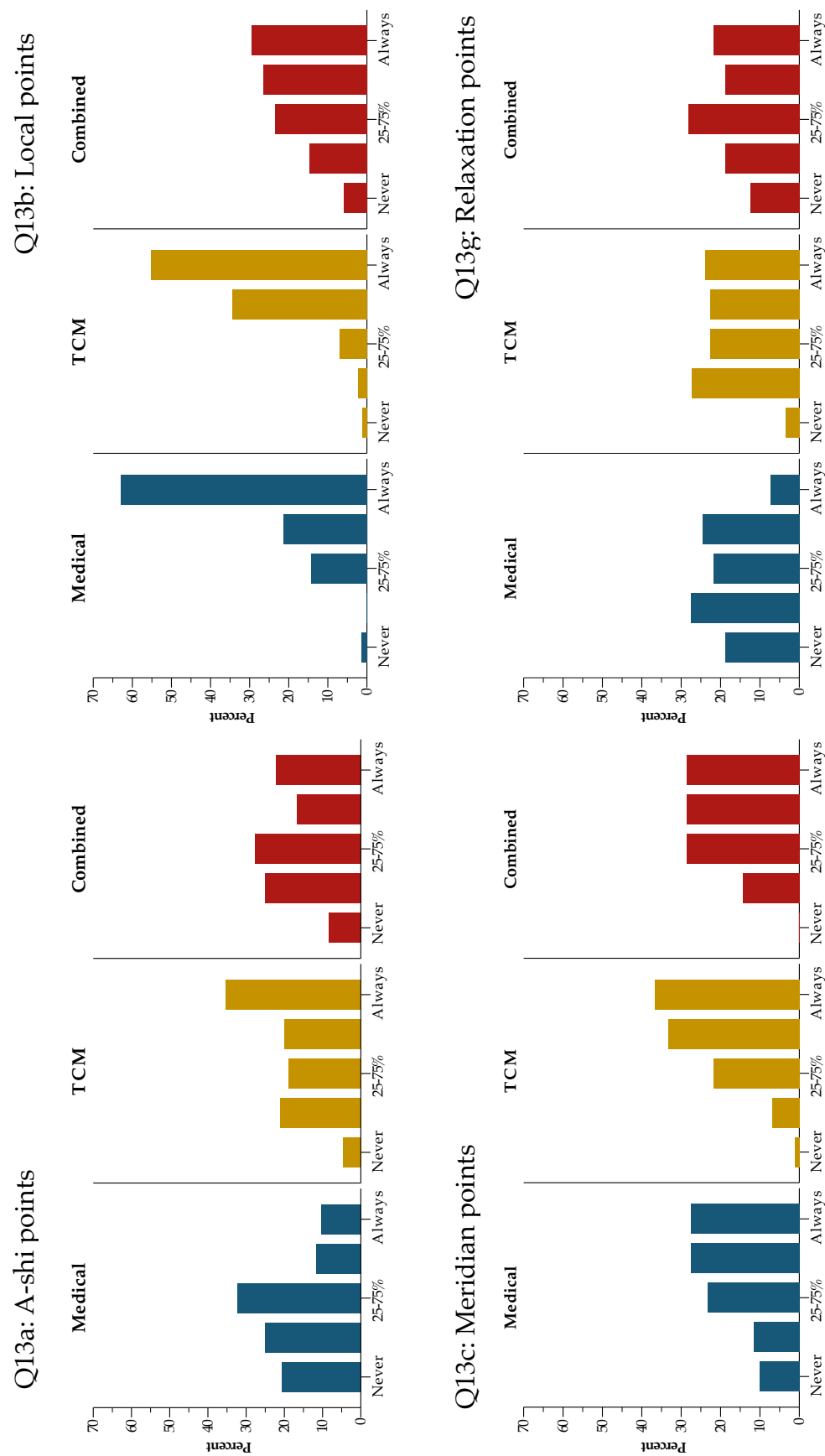
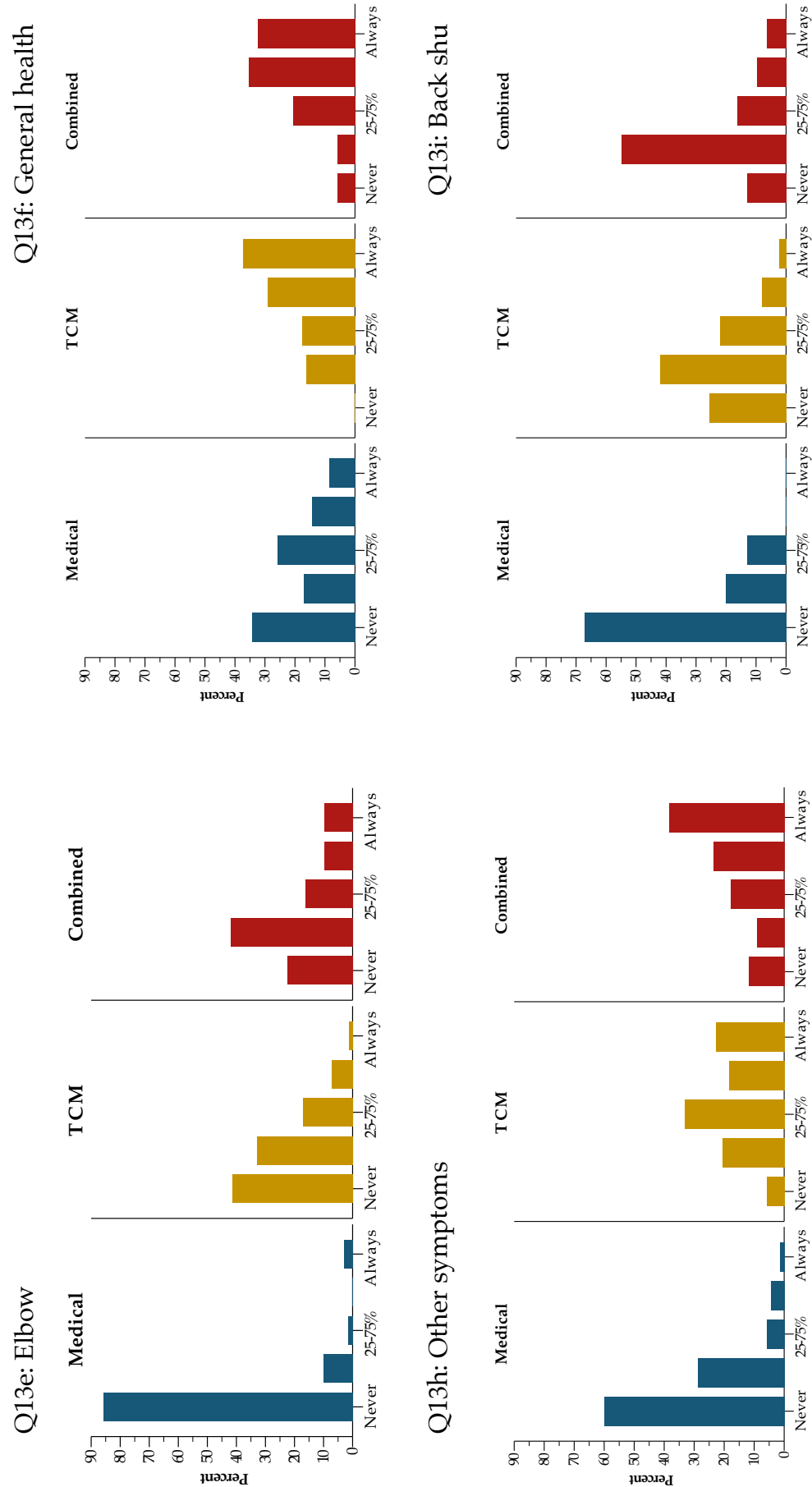


Figure-VI.3: Distribution graphs: D-5,6,8,9 Elbow, General health, Signs & Symptoms, Back shu



General health and Other symptoms (Q13f-h)

Medical acupuncturist use these modes less than both TBSA groups; the Mann Whitney U test showed these differences were statistically significant ([table-VI.9](#)). Moreover the effect sizes were among the largest, for example, for Other symptoms (Q13h) Combined group compared to Medical $r=0.6$ ([table-VI.9](#)). However, 66% of medical practitioners insert needles to strengthen general health, and for some it is a common part of their practice ([figure-VI.3: Q13f](#)): In addition 40% of medical acupuncturists address Other symptoms (Q13h) on at least some occasions, albeit infrequently ([figure-VI.3: Q13h](#)). Both modes showed clear differences between groups, but neither could be considered purely the preserve of TBSA.

Unaffected knee (Q13d)

Needling points on the unaffected knee, whilst being used on occasions by a majority of practitioners (68%), does not constitute a frequently employed method. Only 10% of all respondents indicated that they would treat the unaffected knee on more than 75% of occasions. The Mann Whitney test showed that both TBSA groups were more likely to use these points than the medical acupuncturists ([table-VI.9](#)). However, there are some medical acupuncturists who needle the unaffected knee quite frequently, and TBSA practitioners who never do so ([appendix-E.9.5](#)).

Needles at the elbow (Q13e)

Medical practitioners are the least likely to use this mode of point selection, 85% never doing so. Suggesting that, if a practitioner stated that they used this technique it would be quite likely that they use a TBSA. However, the majority of TBSA practitioners who insert needles at the elbow for the knee pain do so infrequently, less than 25% of the time ([figure-VI.3: Q13e](#)).

Back-shu points (Q13i)

The Mann-Whitney tests showed that there was a statistically significant difference between Medical and both TBSA groups ([table-VI:Q13.i](#)). Nevertheless, a around 30% of medical acupuncturists use back-shu points on

occasions. The majority of TBSA use these points on some occasions but infrequently, with the most common response being 'less than 25% of the time'.

VI.4.4. Stimulation (section B)

This section looks at ways in which the acupuncture points are stimulated. Three questions concern the basic parameters of needling: depth of insertion, gauge and retention time (Q14-16). Other questions investigate the method of stimulation ([table-VI.8](#)) using the stem:

The proportion of chronic knee pain treatment sessions for which

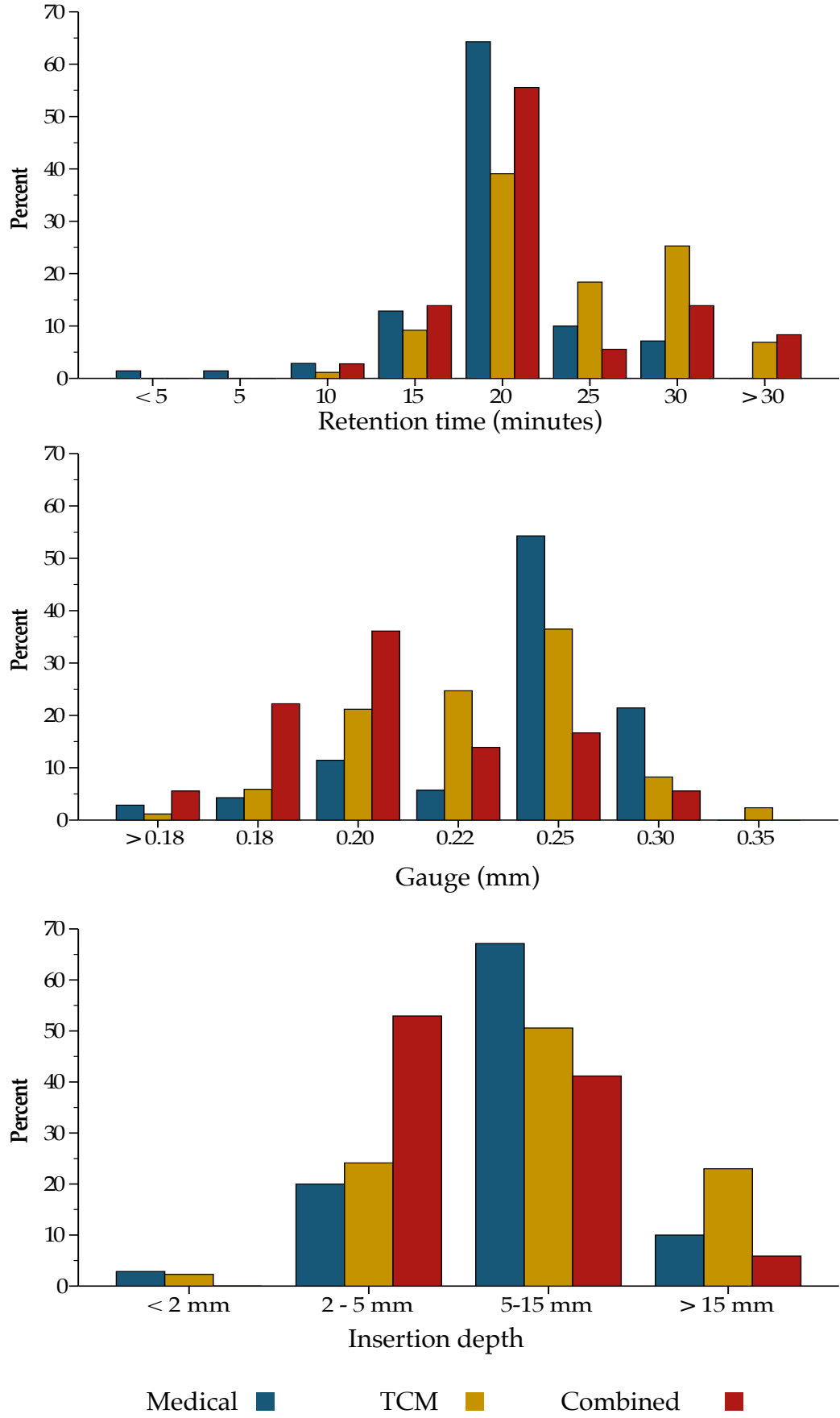
Table-VI.10: Stem Questions - Stimulation Q17

	Question	Abbreviation
17a	I try to obtain a tingling or aching sensations felt by the patient (deqi)	Deqi
17b	I manipulate the needle in order to 'tonify' or 'reduce'	Tonify
17c	I use electro-acupuncture on points local/ adjacent to the affected knee	EA-local
17d	I use electro-acupuncture at distal points on the affect leg	EA-leg distal
17e	I use electro-acupuncture on the opposite leg/ hands/ arms/ torso/ head	EA-general
17f	Intra dermal needles	Intra-dermal

VI.4.4.a. Needling variables (Q14-16)

The Kruskal-Wallis test indicated statistically significant differences for all three questions ([appendix-E.12.1](#)). The most commonly used gauge for both medical and TCM practitioners is 0.25mm, whereas the Combined group generally use thinner, 0.20mm, needles ([figure-VI.4](#)). The follow-up Mann-Whitney showed that the difference between the Combined and both Medical and TCM groups were statistically significant ([table-VI.11](#)).

Figure-VI.4 Needle - depth, gauge, and retention time



The Mann-Whitney tests showed TCM practitioners retain the needles longer than medical practitioners (table-VI.11). However the mode (average) retention time is the same for all three groups, 20mins (figure-VI.4). Overall most practitioners insert the needles to a depth of 5-15mm, although the mode for the Combined group was 2-5mm (figure-VI.4). As can be seen from figure-VI.4 there is heterogeneity within each group. Almost the entire range of options were selected by at least some practitioners from each group for each variable.

Table-VI.11: Stimulation (Q14-16) Mann-Whitney U tests

	Medical vs TCM		Medical vs Combined		TCM vs Combined	
	U	<i>p</i>	U	<i>p</i>	U	<i>p</i>
Needle retention time	1923	< 0.001	1094	0.206	1250	0.064
Needle thickness	2189	< 0.001	563	< 0.001	919	< 0.001
Insertion depth	2831	0.395	848	0.007	1009	0.003

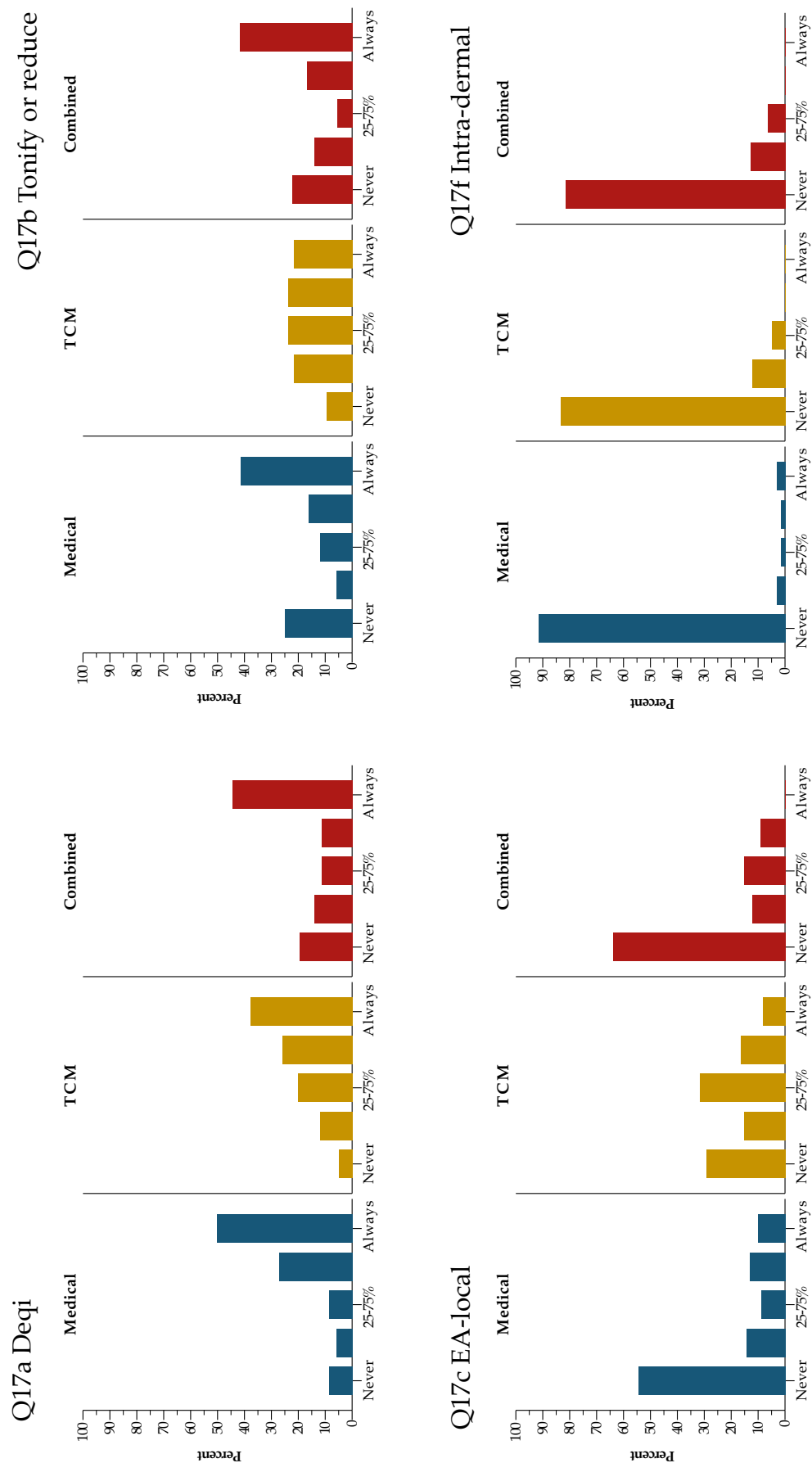
Asymp. Sig. (2-tailed)

Non significant differences shaded $p < 0.05$

VI.4.4.b. Needle manipulation (Q17a-b)

Manipulation of the needles to obtain deqi is common among practitioners of all groups. In each group the majority of practitioners try to obtain deqi on more than 75% of occasions (figure-VI.5: Q17a). Similarly, most practitioners stated they use tonifying and reducing techniques, including medical acupuncturists. Only 25% of medical acupuncturists 'never' tonified or reduced (figure-VI.5: Q17b). The Kruskal-Wallis test detected no statistically significant difference between groups for either of these components.

Figure-VI.5: Stimulation graphs: Q17 (a,b,c,f) Deqi, Tonify, EA-local, Intra-dermal



VI.4.4.c. *Electro-acupuncture (Q17c-e)*

A majority of the TCM (71%) and a minority of the Medical and Combined groups use electro-acupuncture around the knee (46% and 34%, respectively) (figure-VI.5: Q17c). Fewer practitioners use electro-acupuncture at distal points on the affected limb (Q17d), 50% for TCM, 25% Medical and 21% Combined. The proportion drops further for EA at points on the torso and unaffected limbs (Q17e) (appendix-E.9.5). Even for TCM only 27% of practitioners use EA on other limbs and torso. This suggests the focus of electro-acupuncture is similar to the modes of point selection (section-VI.4.1). Namely, the area local to the knee is emphasised.

The Kruskal-Wallis tests indicated that for all three questions regarding electro-acupuncture there were statistically significant differences between groups (appendix-E.12.2). There were no significant differences between the Medical and Combined groups. So for these variables it is the TCM group that is different, as they tended to use electro-acupuncture more (table-VI.12).

Table-VI.12: Electro-acupuncture (Q17c-e) Mann-Whitney U tests

	Question	Medical vs TCM		Medical vs Combined		TCM vs Combined	
		U	p	U	p	U	p
17c	EA-local	2325	.011	1002	0.227	862	0.001
17d	EA-leg distal	2237	.003	1099	0.700	1018	0.007
17e	EA-general	2503	.008	1126	0.697	1226	0.115

Non significant differences shaded $p < 0.05$

VI.4.4.d. *Intra-dermal needles (Q17f)*

Intra-dermal needles are rarely used. Only 14% of all respondents use intra-dermal needles, typically using them less than 25% of the time (figure-VI.5: Q17f). No differences were detected between groups (appendix-E.12.2).

VI.4.5. Stimulation - Moxibustion (section B)

The use of moxibustion appears to be the most distinctive feature of TBSAs compared to medical acupuncture. Only five medical respondents were able to use moxibustion in any of their clinics, of these five, only one actually used moxibustion: moxa stick local to the knee. Therefore, the Medical group has been omitted from this section. Sizeable minorities of TBSA practitioners were also not able to use moxibustion in any of their clinics, 26% for TCM and 17% for the Combined styles. Those practitioners who were able to use moxibustion were asked to answer questions on the methods they employed ([table-VI.13](#)).

The proportion of chronic knee pain treatment sessions for which:

TableVI.13: Stem questions - Moxibustion Q20

Question		Abbreviation
20a	I use moxa local to the knee	Moxa-local
20b	I use moxa at points distal to the knee	Moxa-distal
20c	I use moxa on the end of the needles 'warm needle technique' locally to the knee	Warm needle
20d	I use a moxa stick over the knee	Moxa stick
20e	I use moxa placed directly on the skin 'scarring moxibustion' on local points around the knee	Moxa-skin
20f	I use moxa placed on a slice of ginger on local points around the knee	Moxa-ginger
20g	I use moxibustion (Any)	Moxa (any)

Both TBSA groups tend to use moxibustion locally. If we consider the three options of more frequent usage: always, more than 75%, 25-75%; we can see that moxibustion is most commonly used at local points ([figure-VI.6](#)). Of all respondents, 78% selected one of the more frequent options for local points. For distal points, however, only 45% of all respondents chose one of the more frequent options. The Mann-Whitney test found that the Combined group were more likely to use moxa at distal points: $U=481$, $p=0.028$. This mirrors the

tendency for the Combined group to deliver treatments relatively less focussed on the knee than TCM practitioners ([section-VI.4.3.b](#))

The frequency of use of warm needle acupuncture and moxa stick is fairly evenly spread from less than 25% of treatments to always, for both groups ([figure-VI.6](#)). These are the most common methods for TCM practitioners. Less than 30% on TCM acupuncturists use either moxa on the skin (Q20e) or moxa with ginger (Q20f). Those that do use these methods employ them infrequently ([figure-VI.6](#)). Whereas the majority of the Combined group use moxibustion directly on skin frequently, and over a third regularly use moxa with ginger ([figure-VI.6](#)). Overall the Combined group tend to use moxa more regularly (Q20g). The Mann-Whitney test was significant: $U=364, p=0.001$ ([table-VI.14](#)).

Figure-VI.6: Use of moxibustion distributions (Q20)

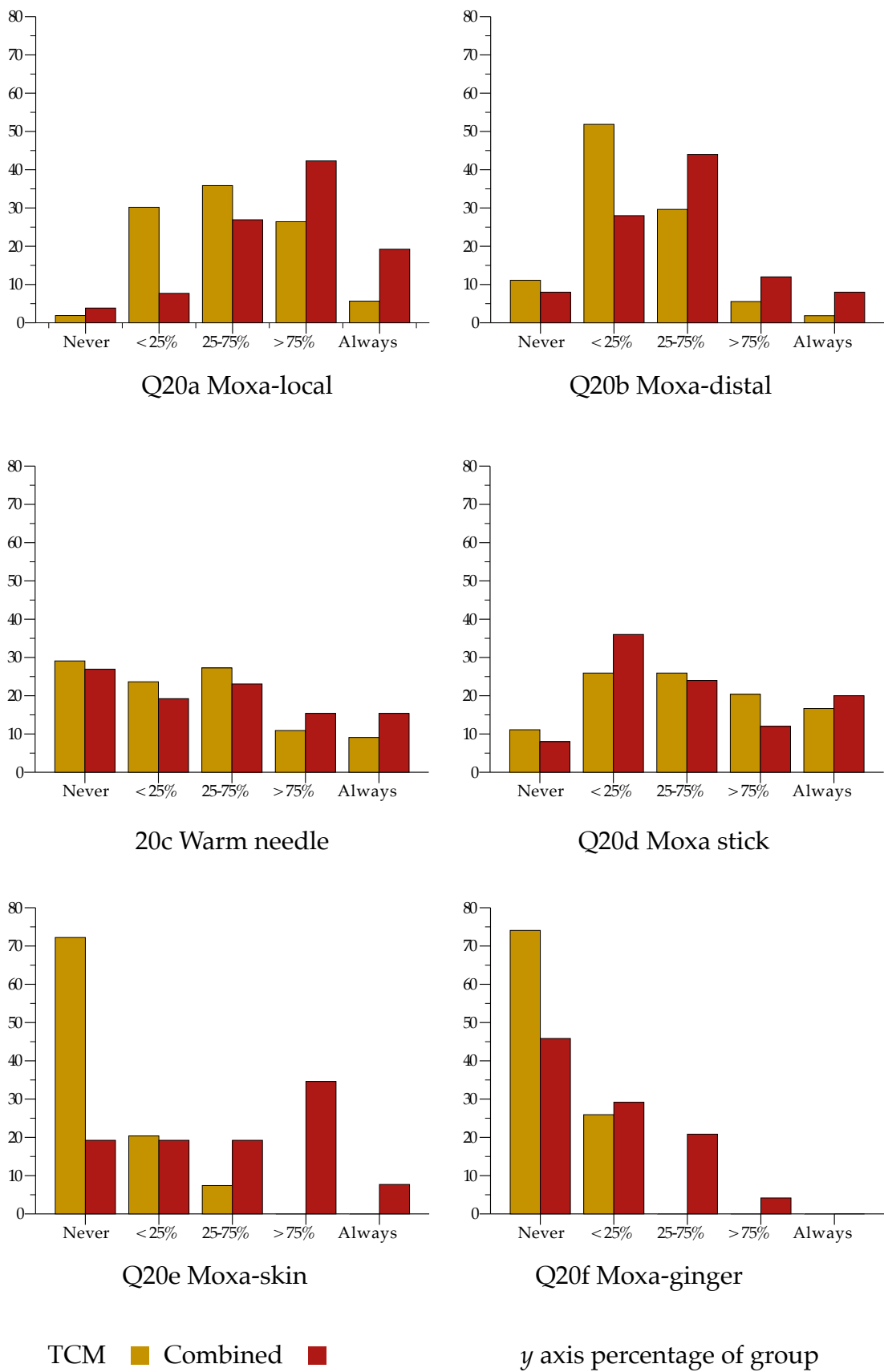


Table-VI.14: Moxibustion (Q20) Mann-Whitney tests

Question		Style	N	Mean Rank	U	<i>p</i>
20a	Moxa-local	TCM	54	39	617	0.240
		Combined	27	45		
		Total	81			
20b	Moxa-distal	TCM	54	36	481	0.028
		Combined	25	48		
		Total	79			
20c	Warm needle	TCM	55	40	645	0.463
		Combined	26	44		
		Total	81			
20d	Moxa stick	TCM	54	40	651	0.795
		Combined	25	39		
		Total	79			
20e	Moxa-skin	TCM	54	32	230	< 0.001
		Combined	26	59		
		Total	80			
20f	Moxa-ginger	TCM	54	35	423	0.004
		Combined	24	49		
		Total	78			
20g	Moxa (any)	TCM	53	34	364	0.001
		Combined	25	51		
		Total	78			

Non significant differences shaded $p < 0.05$

VI.4.6. Course of treatment (section C)

The typical initial course of treatment is five to six sessions (inter-quartile range) for all practitioners. The Kruskal-Wallis test found no significant differences between the groups: $\chi^2(2)=3.8$, $p=0.15$. Practitioners consider discontinuing treatment, if unsuccessful, after between three to five sessions (inter-quartile range). The medical practitioners consider discontinuation approximately one session earlier than the TBSA, and this was statistically significant (table-VI.16).

This section also contains questions regarding the frequency treatment (table-VI.11) using the following stem:

The proportion of chronic knee pain patients for which:

Table-VI.15: Stem questions - Course of treatment Q23

	Question
23a	I treat once a week
23b	I treat 2 times a week or more frequently
23c	I will suggest to the patient that they come back for top up treatments/ or decrease the frequency of treatment to once a fortnight/ months etc

Two thirds of practitioners said that ‘more than 75%’ of their patients were treated once a week ([figure-VI.7](#)). A similar number of practitioners stated that some patients were treated twice or more a week. A majority of practitioners, in all three groups, provide follow-up treatment to more than 75% of patients; Combined 78 %, TCM 59%, Medical 50% ([figure-VI.7](#)). Less than 10% of all respondents never give follow up treatments. A statistically significant difference was detected only for the follow-up treatments ([table-VI.16](#)). This was due to the Combined group tending to provide more patients with follow-up treatments ([figure-VI.7](#)).

Figure-VI.7 Course of treatment graphs

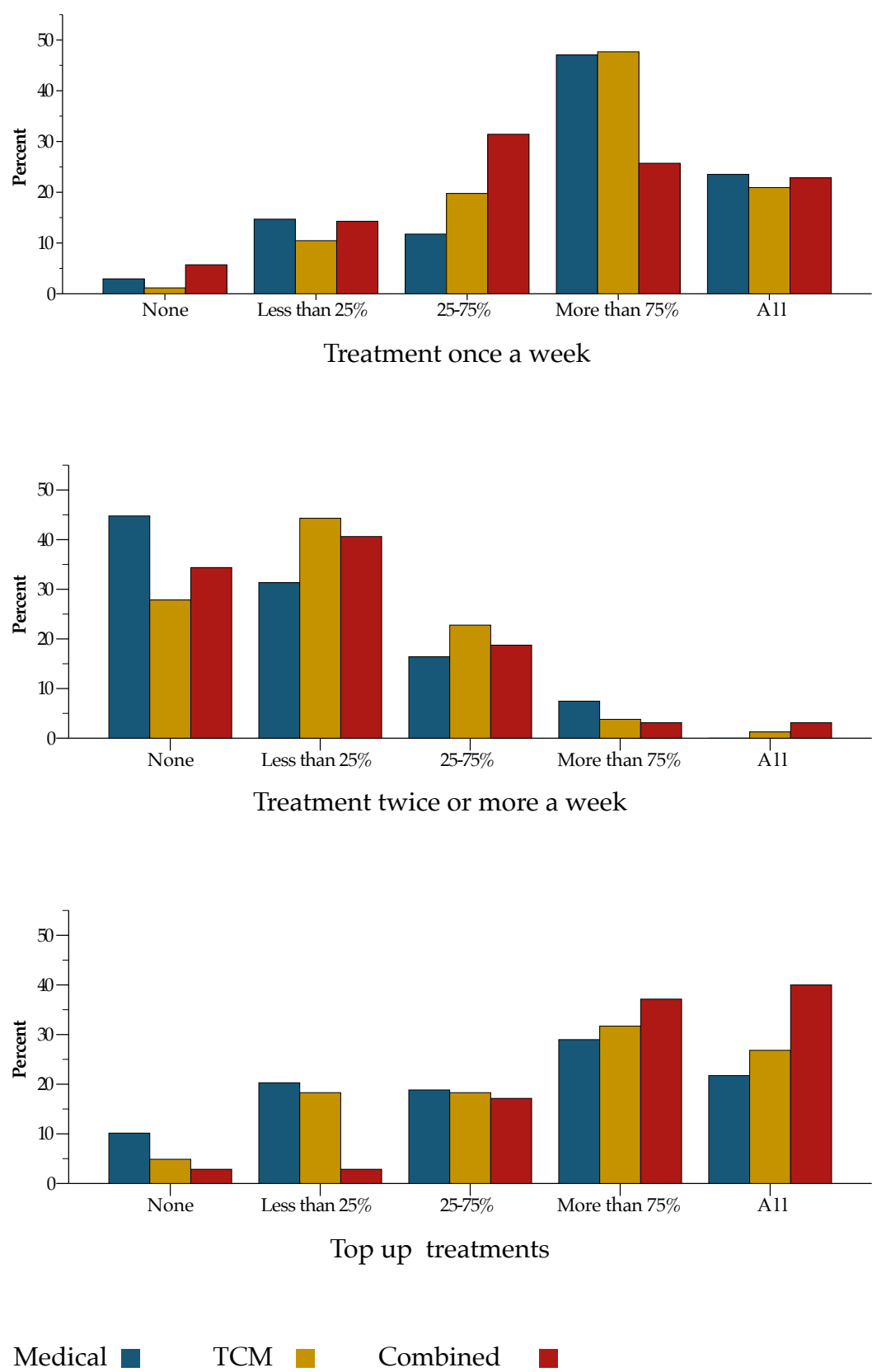


Table-VI.16: Section C Course of treatment Mann Whitney U test

	Medical vs TCM		Medical vs Combined		TCM vs Combined	
	U	<i>p</i>	U	<i>p</i>	U	<i>p</i>
Top up treatments	2524	0.242	800	0.004	1093	0.034
Q22 Discontinuation	1669	< 0.001	782	0.001	1335	0.298

Non significant differences shaded $p < 0.05$

VI.4.7. Adaption of treatment (section D)

This section is concerned with factors that may influence the selection of acupuncture points and amount of treatment given. There are six questions, four address psychological considerations and two relate to time and expense (table-VI.11). The stem is:

The proportion of chronic knee pain patients for which:

Table-VI.17: Stem questions - Adaption of treatment Q24

	Question	Abbreviation
24a	The first one or two treatments will be modified to allow the patient to get used to acupuncture. For example using fewer needles or gentler stimulation	Early treatments
24b	I will reduce the number of needles because I feel the patient is anxious	Anxiety
24c	I will reduce the number of needles because the patient is sensitive to needling	Sensitivity
24d	I will use fewer distal points because I feel the patient believes it is not logical to place needles away from the knee.	Not logical
24e	Finances or time constraints prevent what I believe to be optimal acupuncture being delivered	Time & money
24f	Limited appointment time	Appointment time

VI.4.7.a. Psychological adjustment (Q24a-d)

For three of the psychological factors approximately 90% of practitioners adjust the treatment for some of their patients: to allow patients to get used to acupuncture (Q24a), for anxiety (Q24b), and the sensitivity to needles (Q24c)

(figure-VI.8). The Kruskal Wallis test showed that there are differences between groups for modification in the first treatment (Q24a) and for anxiety (Q24b) (appendix-E.16). The Mann Whitney test found that the Combined group are less likely than the other two groups to make these adaptations, to a statistically significant degree (table-VI.18).

Table-VI.18: Section D: Adaption of treatment Mann Whitney U

	Medical vs TCM		Medical vs Combined		TCM vs Combined	
	U	<i>p</i>	U	<i>p</i>	U	<i>p</i>
Early treatments	2571	0.143	656	< 0.001	950	0.001
Anxiety	2785	0.579	928	0.046	1054	0.009
Time & money	1968	< 0.001	1037	0.198	1206	0.124

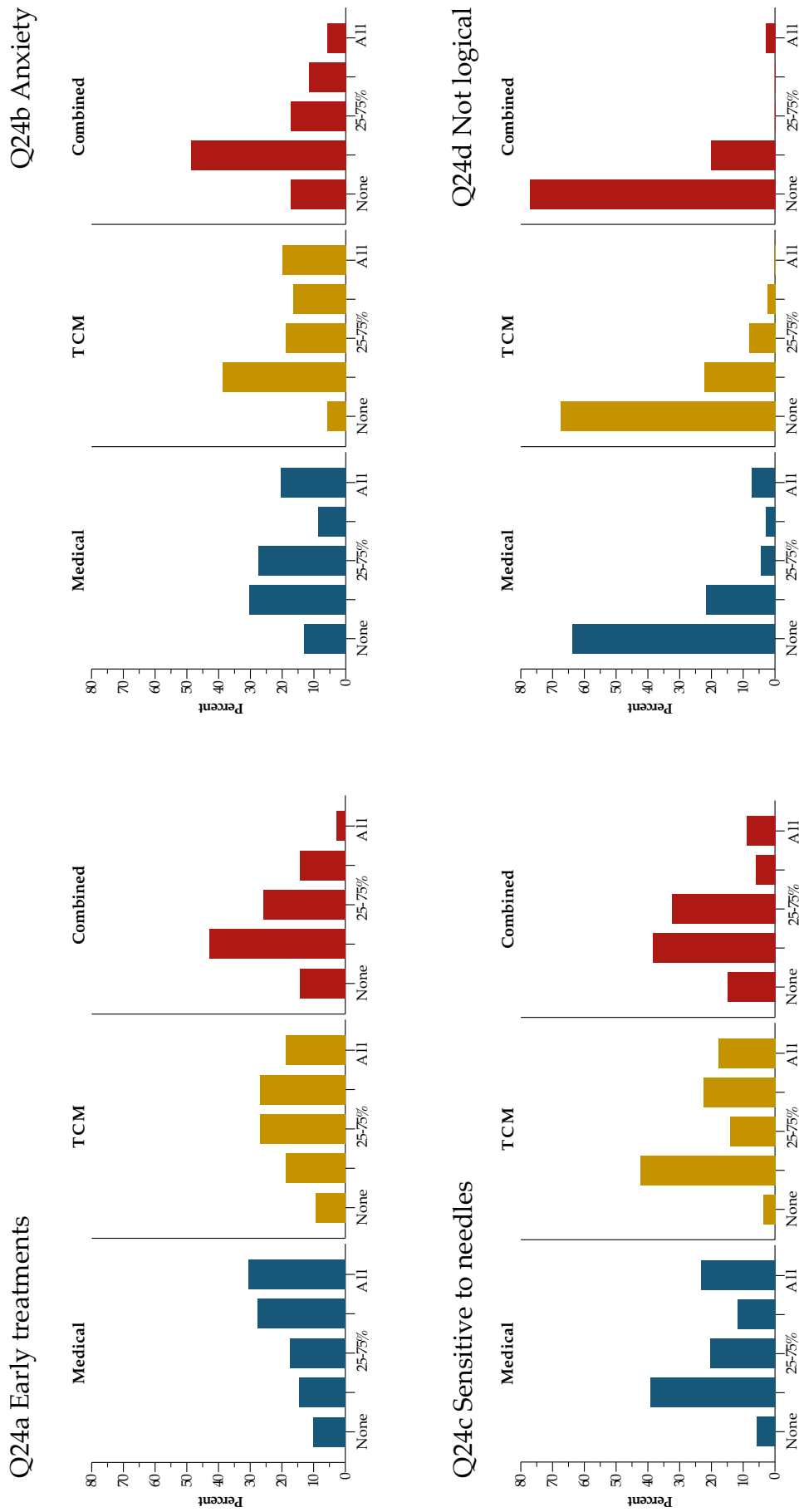
Non significant differences shaded $p < 0.05$

Around 32% of respondents adapt treatment to accommodate a patient who may feel it is not logical to insert needles away from knee (Q24d). Those who might make this adjustment typically do so for less than 25% of patients (figure-VI.8).

VI.4.7.b. *Practical adjustment*

The majority of practitioners (83%) perceive the challenges of expense and time as preventing optimal practice from being delivered for at least some patients. However, it only appears to be an issue for a minority; the most commonly selected option was less than 25% of patients. To a lesser extent, limited appointment time can also be an issue, 50% of practitioners identified this as a problem in some cases. There is a tendency for TBSA practitioners to feel external limitations impact on treatment more commonly than medical acupuncturists (appendix-E.15.5/6).

Figure-VI.8: Treatment modification graphs



VI.4.8. Miscellaneous - Auxiliary methods (section E)

This section looks at the auxiliary methods used (table-VI.13). The stem question is:

The proportion of chronic knee pain treatment sessions for which:

Table-VI.19: Stem questions - Auxillary methods Q25

	Questions	Abbreviation
19a	I use cupping locally to the knee	Cupping-local
19b	I use cupping at places other than locally to the knee	Cupping-distal
19c	I use a heat lamp over the knee	Heat lamp-local
19d	I use a heat lamp at places distal to the knee	Heat lamp-distal
19e	I use massage	Massage
19f	I use TENS	TENS

The auxiliary methods include cupping, heat lamp, massage and TENS. Apart from TENS, the Kruskal-Wallis tests indicated that there are statistically significant differences between groups for each of these auxiliary methods. In every case it is the medical acupuncturists who use these methods less (table-VI.20) (appendix-E.18).

Virtually no medical acupuncturist use cupping or a heat lamp (figure-VI.9). Just over half of TBSA practitioners use a heat lamp and around 40% might use cupping. Very few practitioners frequently use these methods and they tend to be used more commonly over the knee than other places (appendix-E.17.2/4). There was a statistically significant difference between the Medical and both TBSA groups in regards to the use of massage (table-VI.20). Although over 40% of medical acupuncturist use massage (figure-VI.9). TENS is the least used auxiliary method by TBSA practitioners (figure-VI.9).

Table-VI.20: Auxiliary methods Q25 Mann Whitney U tests

	Medical vs TCM		Medical vs Combined		TCM vs Combined	
	U	p	U	p	U	p
Cupping-local	1829	< 0.001	725	< 0.001	1508	0.922
Cupping-distal	1900	< 0.001	722	< 0.001	1502	0.895
Heat lamp-local	1051	< 0.001	683	< 0.001	1134	0.027
Heat lamp-distal	1564	< 0.001	782	< 0.001	1307	0.204
Massage	1474	< 0.001	922	0.033	1164	0.038

Non significant differences shaded $p < 0.05$

VI.4.9. Miscellaneous - Advice (section E)

The survey has three questions regarding the advice practitioners give: diet, exercise, and keeping the knee warm. These use the stem question format “The proportion of chronic knee pain patients for which:” Most practitioners give patients advice on all three areas. The responses, regarding how often advice is given, were evenly spread across the options. Some practitioners always give advice, whilst others do so to less than 25% patients ([figure-VI.10](#)). The Kruskal-Wallis test indicated statistical significant differences between groups ([appendix-E.19](#)). The follow-up pairwise comparisons showed that it is the medical acupuncturist group who differ from the two TBSA groups ([table-VI.21](#)). TBSA practitioners give advice more frequently regarding diet and keeping the knee warm. Medical acupuncturists are more likely to give advice on exercise.

Figure-VI.9: Auxiliary methods graphs

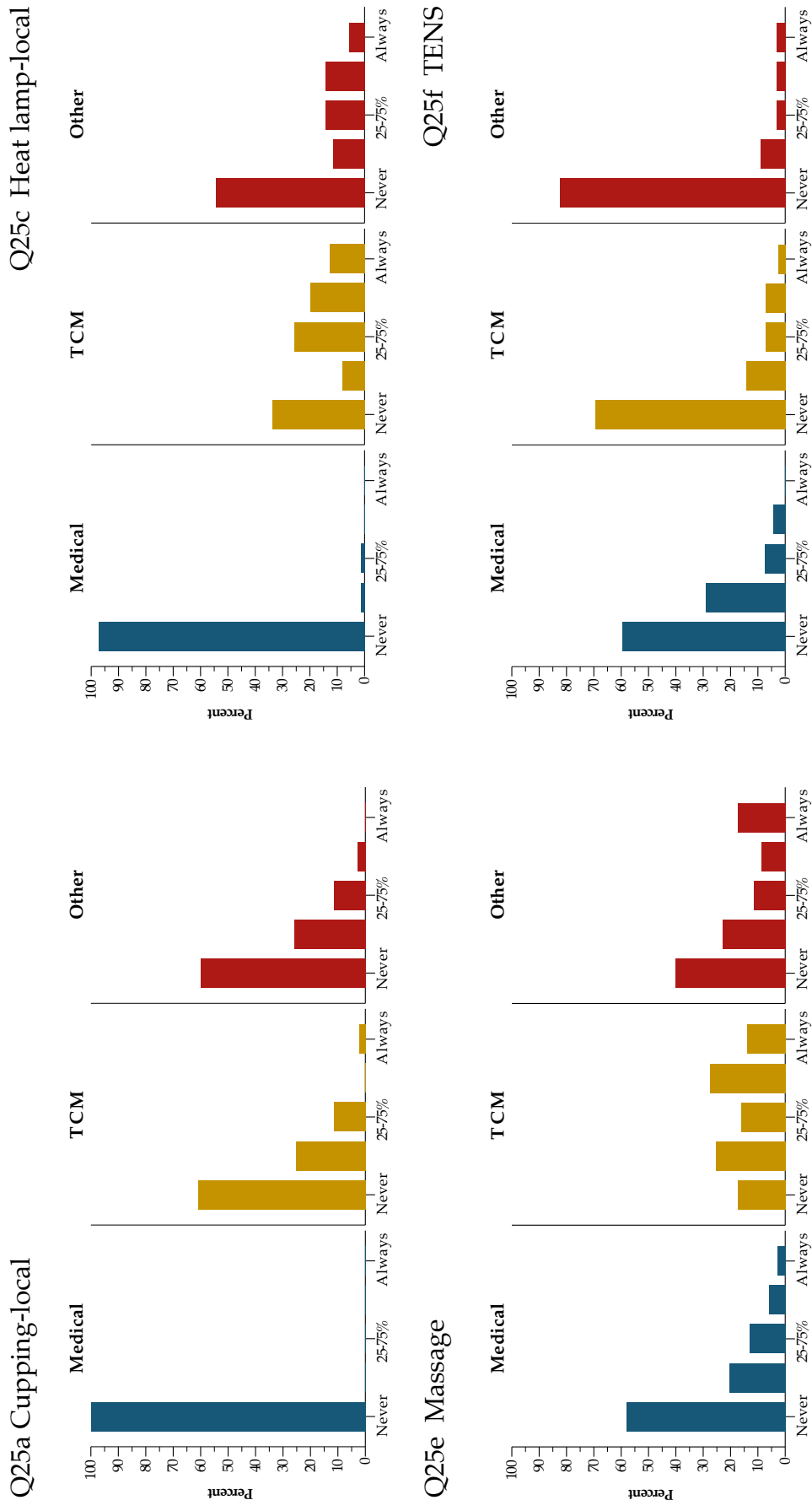


Figure-VI.10: Advice graphs

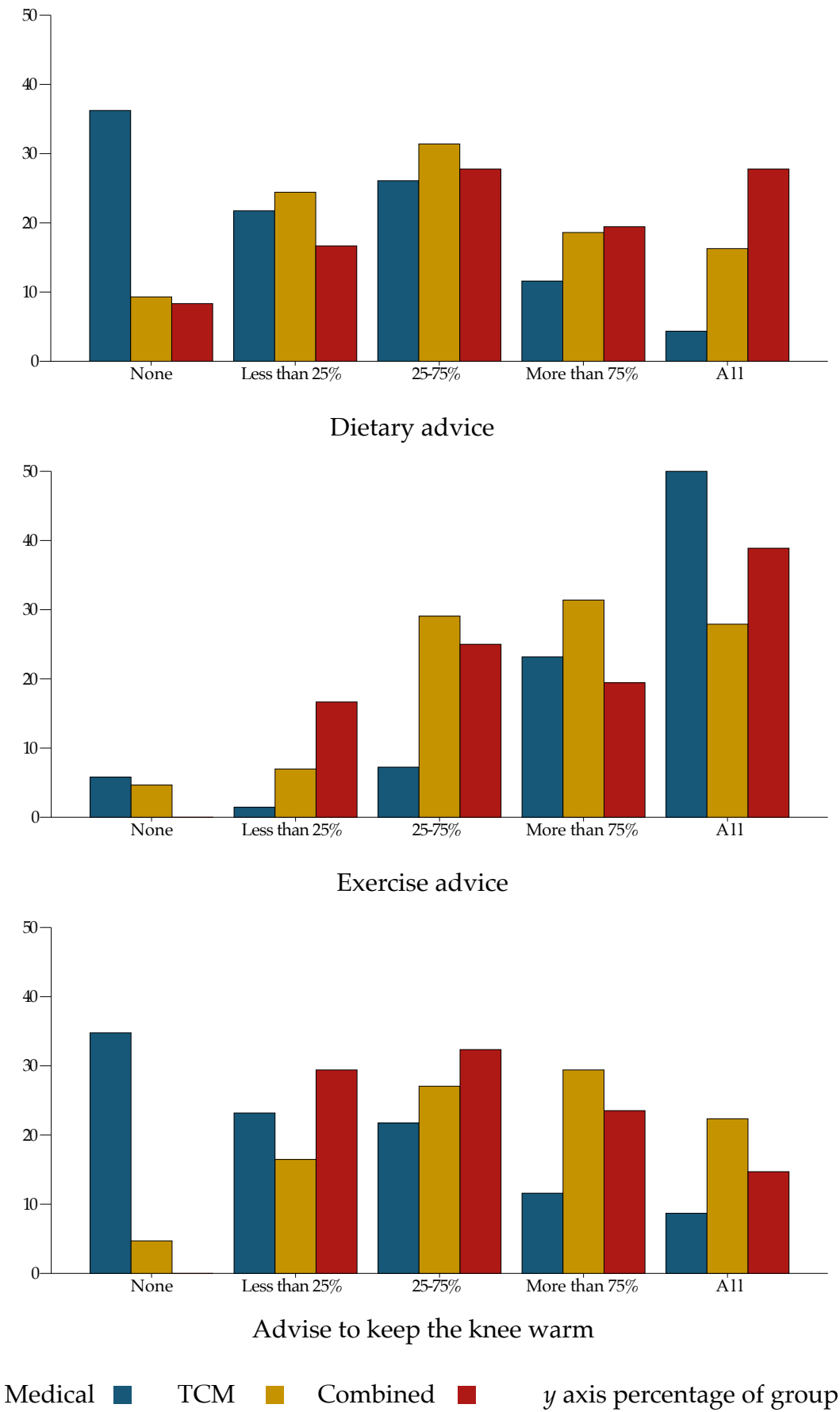


Table-VI.21: Advice Q26 Mann Whitney tests

	Medical vs TCM		Medical vs Combined		TCM vs Combined	
	U	p	U	p	U	p
Dietary advice	1892	< 0.001	653	< 0.001	1310	0.170
Exercise advice	1849	< 0.001	886	0.008	1476	0.674
Advise the knee should be kept warm	1548	< 0.001	702	0.001	1240	0.213

Non significant differences shaded $p < 0.05$

VI.4.10. Safety

Of the 211 respondents, 26 described adverse events, 116 stated there had been no events and 69 left the question blank. Of the 26 who described events the majority indicated bruising. A few respondents noted patients had felt faint or experienced a brief exacerbation of the knee pain. The more serious events are presented in [box-VI.2](#).

Box-VI.2: Adverse events - free text comments

- The patient's knee swelled dramatically the day after treatment, then after a further day returned to its previous size and felt better than it had done prior to treatment'.
- I had one lady who told me she felt sick after I had inserted the needles. I immediately removed them and she then vomited. She had had acupuncture many times before, but told me subsequently that she not been feeling quite 'right' just before treatment on this occasion and she was laid up in the bed the next day with a stomach bug.
- The patient fainted (fortunately on the couch). She was very sensitive to medication/physio etc
- After treatment patient experienced instability in the affected knee that lasted several days, but which did resolve. Only 1 treatment was given. Patient described knee 'giving away' on several occasions following treatment.
- Yes - with over anxious patients I have experienced fainting and 2-3 vaso vagal events.

VI.4.1. Additional methods

Respondents were also given the opportunity to describe any additional methods that were not already covered by the questionnaire. The majority of practitioners completed this section. However, most respondents either emphasised an aspect of treatment they felt was important and/or gave further details. For example, some elaborated on the type of exercise, the exact points used, how to insert the needle. Others reiterated the importance of an integrated approach; 'never just acupuncture'; 'exercise and stretching'. The importance of a full pattern differentiation was also highlighted; 'I give a TCM diagnosis, I treat the patterns of disharmony primarily and the knee as a secondary focus...'. Some additional methods were suggested, including: ear acupuncture, k tape, ice, herbal medicine, supplements, ultrasound, magnets, guasha, orthotics.

VI.5. Discussion

The survey achieved 211 completed questionnaires, making it possibly the largest survey of acupuncture in practice for any condition. An earlier survey into lower back pain received responses from 56 practitioners (Sherman et al., 2001b). Previous research that gathered similar information have been embedded within clinical trials (MacPherson et al., 2004, Sherman et al., 2001a), and examined the practice of relatively few practitioners. In addition, these studies focused on the TCM style. This survey has provided data on a broader range of practitioners than had previously been available.

A variety of methods were used to distribute the survey. Three were electronic means, paper copies posted to practitioners was the fourth method (section-VI. 3.3). The paper copy format nearly reached the anticipated target of 40% response rate. There were 172 electronic responses compared to the anticipated 280. The three electronic methods of distribution used a link to Survey-monkey. Unfortunately, only one link was used for all three electronic methods of

distributing the survey. In hindsight, it would have been better to have created three separate links, this would have enabled data to be collected on whether the respondent had accessed the survey via methods 1,2 or 4 ([section-VI.3.3](#)). The relative success of these methods could then have been evaluated.

The survey data relied upon the recall of practitioners. Responders were not requested to examine their patient notes prior to completing the survey ([appendix-E.2](#)). It is difficult to assess whether relying on recall affected the results in any way.

The ordinal categories for many of the multiple choice questions used percentages rather than adverbs such as 'often' or 'seldom'. It was hoped that this would provide a clearer concept of how frequently each of the components were employed. However, it is possible that some practitioners may have found this format off-putting. This may explain why 19 of the 230 people who started the survey stopped once they have completed the background information ([section-VI.4.1](#)). It is not believed that a slightly different question format would have made any difference to the key findings (below).

All surveys need to balance convenience with the desire to collect greater quantities of data. No information was collected on specific acupuncture point, which had been done in previous research (MacPherson et al., 2004, Sherman et al., 2001a). For most of the modes of point selection, someone familiar with acupuncture should be able to infer which acupuncture points are used. The exception is points selected to address other signs and symptoms (Q13h). It would have been unreasonable to expect practitioners to list of all the potential points they may use to address others signs and symptoms. Such a list could include every acupuncture point, and therefore would be meaningless.

The survey has shown there are differences between styles of practice, in terms of their physical components. However, most of these can only be considered general tendencies rather than defining characteristics. This is because the inter-

style differences are outweighed by the intra-style heterogeneity. This pattern is repeated in almost every section of the survey.

As point selection is supposed to be driven by different theoretical systems, it might be presumed that point selection would be one of the defining characteristics of particular style. Moreover, there would be clear differences between Medical acupuncture and TBSAs. This does not appear to be the case. Medical acupuncturist use most of the modes of point selection less than their TBSA counterparts ([table-VI.8-9](#)). However, with the exception of points at the elbow (Q13e), each mode is used on occasions by at least a third of medical acupuncturists. Points at the elbow are rarely selected by the majority of TBSA practitioners. Consequently, none of the modes clearly delineate TBSA from medical acupuncture. Interestingly, for the most commonly used mode by TCM practitioners, local points (Q13b), there is greater similarity between TCM and medical acupuncturists, than between TCM and the Combined group.

The simple labels of TCM, Medical or other styles of practice may potentially be misleading. One cannot assume that these systems lead to different acupuncture points being selected. In fact, the survey data suggests that a single description can illustrate the typical treatment for practitioners in all three groups: a total of eight needles, five of which are placed around the knee with three meridian and/or general health points. Some TCM and medical acupuncturists frequently provide virtually identical treatments, yet, two practitioners who both consider their style to be TCM may deliver radically different treatments. This is most obvious when considering the number of needles inserted; as some TCM practitioners never insert more than four needles, whilst others never insert less than 20 ([table-VI.6](#)). The pattern is repeated in the first part of section B, stimulation methods. Inter-group differences are detected for variables such as needle retention time ([table-VI.11](#)). However, the mode for all three groups was 20 minutes. This indicates that the inter-group differences are derived from the intra-group heterogeneity. Some

medical acupuncturist retain the needles for particularly short periods of time, and more TCM practitioners leave the needles for slightly longer.

The survey asked practitioners to identify their own style of practice, due to relatively small numbers the Japanese and Five-element practitioners were allocated to a single Combined group. If sufficient responses had been obtained from these practitioners a more homogenous picture of these styles may have been identified. However, most of these practitioners were also trained in TCM and may, like the experts interviewed ([section-V.5.4.d](#)), alternate styles. The responses of the medical acupuncture group, which contained only those practitioners who had identified themselves as purely medical acupuncturists, suggested that TBSA ideas are used to a greater degree than might be anticipated by studying the medical acupuncture literature ([section-III.6.3](#)).

When conducting systematic reviews, the description of protocols as being based on a particular style of practice will not be enough to ensure that different trials are sufficiently homogenous. More detailed methods of evaluating the similarity between acupuncture protocols will need to be developed.

The use of moxibustion is a defining feature between medical acupuncture and TBSAs. Effectively, medical acupuncturists do not use moxibustion. The Combined group tend to use a wider range of moxibustion techniques than TCM practitioners. The use of moxibustion also echoed the finding of the distribution of acupuncture points, the TCM group focus slightly more of the area local to the knee ([table-VI.11](#)). The auxiliary methods also appear to be features that can delineate medical acupuncture from TBSAs. Almost no medical acupuncturist use cupping or a heat-lamp. It is surprising that so few medical acupuncturists use a heat lamp. It is also somewhat ironic that an electric heat lamp can be considered part of “traditional” rather than medical acupuncture.

The survey showed that a large majority of practitioners adapt the treatment to accommodate patients who were anxious, sensitive, or new to acupuncture.

This has important implications for clinical research. Protocols should be designed with the capacity to allow for some moderation in the treatment should participants be particularly anxious or unaccustomed to acupuncture.

The survey focuses on the physical process, therefore, the clinical reasoning and beliefs of practitioners were not explored. A number of questions emerged from the data that point to the need for future qualitative research. Medical acupuncturists are less likely to use a-shi points than TCM practitioners. This is surprising as it had been thought that medical acupuncturists concentrate on treating local tender points, because of the known potential neurological mechanisms (Bishop et al., 2011c, Bradnam, 2011). The medical acupuncturists who were interviewed also indicated that they use the 'traditional anatomical locations' ([section-V.5.1.e](#)). This suggests medical acupuncturists place greater emphasis on traditional descriptions of point locations than the TBSA practitioners. Further research needs to be conducted to explore beliefs regarding the nature of acupuncture points.

Another intriguing finding is the number of medical acupuncturists who use tonifying and reducing techniques, only 25% indicated that they never did so ([figure-VI.5: Q17b](#)). Medical acupuncture has been defined as using existing biomedical knowledge (White and Editorial-Board, 2009). The word 'tonifying' is defined as a concept of Chinese medicine (Oxford Living Dictionaries, 2018). Whilst potential mechanisms of the possible effects of these techniques are in the literature (Birch, 2009), it is unclear whether this knowledge underpins the use of tonifying techniques by medical acupuncturists.

The Combined group do not adapt their treatments as much as the other two groups. Medical practitioners might see a higher proportion of patients who are new to acupuncture. However, this would not explain why TCM practitioners also adapt the treatment more frequently. Practitioners in the Combined group may place greater value on delivering the treatment indicated by their

diagnostic theory, and to be less inclined to make adjustments. Further research is required to explore the clinical reasoning behind these adjustments.

Commonly patients are treated once a week rather than more frequently. As the survey did not include a follow up question, it is not clear why this is the case. It could be due to financial or time constraints, or considered the clinically most effective pattern. Further research is required to shed light on this issue.

Most practitioners give patients exercise advice. The level of detail of the advice is not clear, as the question did not specify (Q26). It could range from advice tailored to the individual after clinical examination, or be general encouragement to exercise. This could be explored in future research. It would appear that comparisons of acupuncture versus exercise (Corbett et al., 2013), are not real world scenarios, as in practice acupuncture is frequently combined with exercise advice.

VI.5.1. Summary

There are differences between styles of practice, however, intra-style heterogeneity is much greater. Therefore, many of the inter-style differences can only be considered as broad generalisations. At the level of individual practitioners it would be difficult to identify their style of practice by examining the acupuncture points used. Medical acupuncture is not clearly different from TBSAs, in terms of point selection and stimulation. The delineation between medical acupuncture and TBSAs appears to be the use of moxibustion, cupping and heat lamps. The survey raises a number of questions regarding the clinical reasoning processes of acupuncturists, that should be explored using qualitative methods.

VII. Translational analysis: acupuncture for
osteoarthritis of the knee

转
译
医
学

VII.1. Introduction

A translational research approach is used in this chapter to examine acupuncture for the treatment of osteoarthritis of the knee. It draws on the new knowledge developed in the first three phases of this project. The main focus is a re-evaluation of clinical research. Trials designed to test acupuncture versus placebo have a privileged status among stakeholders and researchers ([section-II.3](#)). Within this thesis it is argued that such a simple dichotomy is not possible.

Placebo is not a meaningful concept in acupuncture research ([section-I.4](#)). Well-designed trials can balance psychological factors between groups, and thereby test specific physical variables of acupuncture. In other words, demonstrate acupuncture is not simply a placebo ([section-I.4.4](#)). However, the controls procedures used - sham acupuncture - are not inert. Therefore, the risk of bias of these procedures needs to be assessed ([section-II.5.4](#)). Fixed or semi-flexible acupuncture protocols, no matter how well designed, will always compromise external validity. Consequently, it is also necessary to evaluate the external validity of the acupuncture protocols used in clinical trials ([section-II.5.4](#)).

This chapter presents a novel method of assessing both experimental and control procedures. Data derived from the practitioner survey is used to examine the generalisability of the current evidence base. Aside from clinical research, acupuncture in practice is compared and contrasted with the theoretical approaches. Avenues of potential future research regarding the clinical reasoning of acupuncturists are identified.

Box-VII.1: Objectives of the translational analysis

- ▶ Evaluate the external validity of the acupuncture provided in clinical trials
- ▶ Evaluate the risk of bias of sham acupuncture procedures
- ▶ Evaluate the generalisability of the evidence base
- ▶ Compare and contrast practice with theory

VII.2. Assessment Methods

The research framework stipulates that clinical research can only be effectively evaluated using a translational approach ([section-II.4.6.b](#)). Two of the axioms the framework are: external validity will always be compromised when using fixed or semi-flexible protocols; sham procedures inevitably retain some components of acupuncture ([box-II.1](#)). The assessment methods for both acupuncture and sham procedures are closely related. The essential principles of which are: an understanding of which components have been omitted or restricted should be used to evaluate external validity; an understanding of which components have been retained should be used to evaluate the risk of bias of sham procedures ([box-II.1](#)).

VII.2.1. Previous assessment methods

The quality of acupuncture in clinical trials for osteoarthritis of the knee has been assessed in previous studies (Ezzo et al., 2001, Purepong et al., 2012). Purepong et al (2012) base their criteria on the STRICTA reporting guidelines (MacPherson et al., 2010); Ezzo et al (2001) employ the recommendations of Stux and Birch (2001). The criteria these studies used are set out in [table-VII.1](#). Both studies reviewed the literature to identify the relevant acupuncture points and benchmarks for other criteria. Purepong et al (2012) also conducted a survey of 11 practitioners. Ezzo et al (2001) used acupuncturists, blinded to the publication and results of studies, to assess the adequacy of treatment. The assessment method applied in this chapter further develops the criteria of two previous studies.

Table-VII.1: External validity assessment criteria - previous studies

Criteria	Ezzo et al (2001)	Purepong et al (2012)
Adequacy	Acupuncturists	X
Points used	Textbooks	Textbooks and survey of 11 practitioners
Number of acupuncture points	8 points per knee	X
Total number of sessions	> 10	10-15
Frequency of sessions	X	2/week
Needle retention time	X	20 minutes
Stimulation of deqi	Yes	X
Electro-acupuncture	high or low frequency	X

VII.2.2. Establishing the criteria and benchmarks

Twenty two criteria have been identified to evaluate acupuncture and sham acupuncture ([table-VII.2](#)). These criteria are based on the components of acupuncture established through the literature review ([chapter-III](#)) and the expert interviews ([chapter-V](#)). Data from the practitioners' survey was also used to estimate the importance of the criteria in practice and establish benchmark parameters.

The median scores of the survey responses are used to evaluate the importance of most of the criteria. Many of the questions required ordinal category responses: never, less than 25%, 25 to 75%, more than 75%, always ([chapter-IV](#)). Each category was labelled numerically 0-4: where never = zero, less than 25% = 1, and so on, to always = 4. Thus, the median score indicates the use of each criterion in practice. For example, a median of 3 for criterion X shows that more than 50% of practitioners use criterion X 'more than 75% of the time'. If a clinical study does not include criterion X, then the external validity is compromised to a relatively large degree. Conversely, if criterion Y has a median of 1, then more than 50% of acupuncturist use criterion Y 'less than 25% of the time'. Therefore, the omission of criterion Y has less impact on external validity.

The survey data was also used to establish benchmark parameters for other criteria. The primary benchmarks for the number of points used locally and overall (Nō 3/22) were obtained using the interquartile ranges. The 10-90th percentile ranges have also been reported, in brackets, as secondary benchmarks ([table-VII.2](#)). The survey collected data on the fewest and the most needles practitioners may use, which showed that the potential ranges were wider ([table-VI.7](#)). However, responses regarding the average number of needles used are more relevant when assessing external validity. Benchmarks were established for needle depth of insertion, needle gauge and retention time (Nō 14-16). For these three criteria the mode (average) response has been used as the primary benchmark. Significant minorities of practitioners use different thicknesses and depths ([figure-VI.6](#)), therefore secondary benchmarks have been provided. The primary and secondary benchmarks capture more than 90% of practitioners. The medians and the benchmark parameters were calculated for all three groups, Medical, TCM and Combined. This enables judgements to be made about which style of practice individual studies reflect.

The practitioner survey ([chapter-VI](#)) did not directly ask how many treatment sessions patients normally received (Nō 17). Previous research identified 10 sessions as the minimum required (Ezzo et al., 2001, Purepong et al., 2012). Therefore, 10 sessions was adopted as the primary cut-off point for adequate treatment. The survey did ask the number of sessions given in the initial course ([appendix-E.7 Q:21](#)). The secondary benchmark was created using the median response to this question.

The criteria also include follow-up sessions (Nō 18), where patients come back for top-up treatment over an extended period of time. The frequency of follow-up treatments is typically decided by patient and practitioner to suit individual circumstances. It is difficult to clearly define what constitutes follow-up treatments in clinical trials, this issue is discussed below.

Table-VII.2: External validity assessment criteria

Nō	Mode	TCM	Medical	Combined
1	Practitioner	x	x	x
2	Local points	4	4	3
3	Local points [†]	4-6 (3-7)	4-6 (3-7)	3-6 (1-7)
4	Meridian	3	3	3
5	Help general health	3	1	3
6	A-shi	3	2	2
7	Relaxation	2	2	2
8	Other S&S	2	0	3
9	Unaffected knee	1	1	1
10	Back Shu points	1	0	1
11	At the elbow	1	0	1
12	Reduce needles	2	3	1
13	Deqi	3	3.5	3
14	Depth of insertion ^{††}	5-15mm (> 2)	5-15mm (> 2)	2-5mm (> 2)
15	Needle thickness ^{††}	= 0.25mm (> 0.2)	= 0.25mm (> 0.2)	= 0.2mm (> 0.2)
16	Insertion time ^{††}	20 mins (>15)	20 mins (>15)	20 mins (>15)
17	Session	> 10 (≥ 6)	> 10 (≥ 6)	> 10 (≥ 6)
18	Follow-up sessions	3	3	3
19	Moxibustion	2	0	3
20	Electro-acupuncture	2	0	0
21	Cupping	0	0	0
22	Total points [†]	7-12 (5-16)	5-8 (4-10)	5-9 (4-10)

[†] interquartile range (10-90 percentile)

^{††} mode (> 90% of practitioners)

The survey included a number of questions regarding adapting treatment because the patient was sensitive, anxious, new to acupuncture, or would find the use of points away from the knee illogical ([section-VI.5.7.a](#)). These have been conflated into one criterion Nō 12 - Reduce needles. If a protocol allowed a flexible number of needles, then it seems likely that practitioners would naturally modify the treatment for any of the reasons identified in [section-VI.5.7.a](#). The median response to the question ‘reducing the number needles in the first few treatments’ ([table-VI.17](#)) was used to establish the benchmark for Nō 12. This was the most common reason given for reducing the number of needles ([figure-VI.8](#)).

Practitioners may also choose not to strengthen the deqi sensations. A criterion to assess whether studies allowed needle stimulation to be minimised is not included in this evaluation as this information is not reported in any of the studies. Probably, if a patient was uncomfortable the practitioner would not continue to manipulate the needle.

VII.2.3. Methods of evaluating each criterion

To evaluate external validity, each criterion is assessed using a colour coding system - red, amber, green and grey. Red indicates that the protocol does not include the criterion, therefore does not have complete external validity. Amber indicates that the criterion is included, but there are some limitations. Green indicates that the criterion is similar to practice. Grey is used when there is either not enough information or the criterion is not applicable. When there is ambiguity regarding these judgements explanations have been provided (tables-VII.3-10).

The *modes of point selection* is a concept developed as part of this project (section-III.5.5.a). Naturally, the trial reports do not explicitly describe 'modes of point selection'. The researcher used his knowledge of Chinese medicine, and referred to textbooks (section-III.3.1), to assess which modes were used; based on the acupuncture points reported. For the most part this was straightforward, as the modes are defined through anatomical location. Some points could be considered as indicated for more than one mode. For example SP-6 could be considered a meridian point (Nō 4), a point for general health (Nō 5), or even a point for relaxation (Nō 7). Other issues could also be open to debate, such as the number of potential meridian points required for a protocol to be considered to have fulfilled that criterion. When there was ambiguity, further explanation has been provided within the tables. The local points have been considered to be simply local points. In other words, ST-36 has not been classed as a point for general health, (Nō 5); GB-34 has not been counted as a point for relaxation (Nō 7).

A similar colour coding system is also used to evaluate the sham procedures. Red indicates when the criterion is similar to practice, which implies that it may be active and represents a risk of bias. On the other hand, green indicates that the criterion is not present; therefore meets the intention of a sham control. Amber indicates an intermediate or unknown risk. Grey is used when not enough information was reported or when not applicable. If the real acupuncture does not use a particular criterion, then the sham procedure is assessed as 'not applicable'. For example: if the acupuncture protocol does not use moxibustion then this is judged as red, because moxibustion is part of practice. The sham procedure would then be classed as 'not applicable', for moxibustion. The ideal scenario would be that all criteria to be judged as green for both real and sham procedures, however, this is not possible ([chapters-I&II](#)).

VII.2.4. Selection of studies

The systematic review by Vickers et al (2012) has been identified as the highest quality evidence, to date ([section-IV.2](#)). Therefore, the trials included within this review were selected to represent the current evidence base. Of the 10 trials identified by Vickers et al (2012) four trials used semi-flexible protocols (Foster et al., 2007, Scharf et al., 2006, Williamson et al., 2007, Witt et al., 2005) and four trials used fixed protocols (Berman et al., 2004, Suarez-Almazor et al., 2010, Tukmach et al., 2004, Vas et al., 2004). The remaining two trials used individualised treatments (Lansdown et al., 2009, Witt et al., 2006). Vickers et al (2017) recently published an update to this review which includes an additional two relevant studies; one semi-flexible protocol compared to sham laser (Hinman et al., 2014), the other fixed protocol vs non-insertion (Mavrommatis et al., 2012).

In this translational analysis the external validity of the 10 trials that used either fixed or semi-flexible protocols are evaluated. The style of acupuncture is not always clearly stated but eight of the trials indicate that TCM acupuncture was used; Williamson et al (2007) state that they used Western [medical] acupuncture, Hinman et al (2014) describe their style as a combination of

Western [medical] and TCM acupuncture. For consistency all studies are compared to the benchmarks TCM acupuncture ([table-VII.2](#)). All of the trials described methods of protocol development including literature reviews, panel of experts, surveys, and previous clinical trials, apart from two (Hinman et al., 2012, Hinman et al., 2014, Mavrommatis et al., 2012)

VII.3. Re-evaluation of the evidence-base

VII.3.1. Assessment of external validity

The five semi-flexible protocols have different degrees of flexibility. Foster et al (2007) is the most restrictive with potential points limited to local (Nō-2), meridian (Nō-3) and relaxation (Nō-7) modes of point selection (Hay et al., 2004). There is some capacity to treat general health (Nō-5) ([table-VII.3](#)). Scharf et al (2006) allow more flexibility, with the possibility of selecting points at the elbow (Nō-11), back-shu (Nō-10) and for other signs and symptoms (Nō-9) (Streitberger et al., 2004). Although for each of these modes the actual choice of points is limited ([table-VII.4](#)). Witt et al (2005) have a high degree of external validity in terms of modes of point selection ([table-VII.5](#)). It is only limited by mandating against reducing the number of needles, a minimum of six needles had to be inserted locally to the knee (Brinkhaus et al., 2003). Whilst this is within the interquartile range ([table-VI.5](#)) the majority of acupuncturists will sometimes insert fewer ([appendix-E.9.1](#)). The protocol for Hinman et al (2014) states that ‘other points could be used at the GPs discretion’ (Hinman et al., 2012, p6) - this creates ambiguity regarding the extent to which the standardised points list was used.

The overall number of points, thickness of the needles, retention time, depth of insertion and elicitation of deqi are generally consistent with typical practice for TCM practitioners. Suarez-Almazor et al (2010) used thinner needles and overall fewer points ([table-VII.9](#)). Williamson et al (2007) prescribed predominantly local points, with an option for up to three additional needles.

Therefore in this study, Meridian and General health (Nō-4/5) have been rated amber, as these modes could be selected but relatively few needles created a limitation. Williamson et al (2007) state that they used western acupuncture (equivalent to medical), therefore it has been assumed that other signs and symptoms (Nō-8), back-shu (Nō-10), and at the elbow (Nō-11) modes were not used ([table-VII.6](#)).

The most significant mode of point selection that differs between the fixed protocols is meridian points (Nō-4). There are six meridians that pass through the knee. If a fixed protocol includes one distal point on each meridian, then the affected meridian will be stimulated, regardless of where the pain manifests. Berman et al (2004) stimulated five meridians, Vas et al (2004) and Mavrommatis et al (2012) stimulated four ([tables-VII.7/8/12](#)). They all used SP-6 which is thought to be the junction of three meridians Spleen, Liver and Kidney. Therefore, it could be considered that Berman et al (2004) stimulated all but the Stomach meridian, Mavrommatis et al (2012) and Vas et al (2004) all but the Gallbladder and Bladder meridians. It is difficult to say which has the most external validity. Some practitioners may feel that Berman (2004) is actually weaker, on the basis that the Stomach meridian is particularly important for knee pain. All three studies have been rated amber for Nō-4. The number of meridians stimulated by Suarez-Almazor et al (2010) and Tukmachi et al (2004) were more limited. Suarez-Almazor et al (2010) only used SP-6, this accounts for the three Yin meridians. However there were no points on the Yang meridians, which are more commonly used in painful conditions (Cheng, 1999, Legge and Vance, 1997). Tukmachi et al (2004) only stimulated the Liver and Bladder meridians. These two studies have been categorised as red ([tables-VII.9/10](#)).

One of the semi-flexible protocols, Scharf et al (2006), is also slightly limited for the meridian points (Nō-4). Scharf et al (2006) does not include a distal Gallbladder meridian point; therefore it has been rated amber ([table-VII.4](#)). Otherwise, the semi-flexible protocols are considered to have generally better

external validity for this component (Nō-4) than the fixed protocols. The five fixed protocols have also been judged to have less external validity than the semi-flexible ones because they lack the capacity to treat other signs and symptoms (Nō-8) or use a-shi points (Nō-6). Even if a protocol does not stipulate a-shi points, this does not necessarily mean they were not used. This is because individual practitioners may locate points using palpation. In other words, the practitioner palpates around the immediate vicinity of a pre-specified point, and then inserts the needle where the patient reports the most discomfort.

Only Hinman et al (2014) included the use of back-shu points (Nō-10) (Hinman et al., 2012). Although it is not clear how these points could have been used at the same time as the other points on the standardised list, which are on the anterior aspect of the body. Treatment was given with the patient lying down (Hinman et al., 2012). It is possible that needles were removed and then the second group of points inserted within one treatment session. This would have reduced the length of time each acupuncture point was stimulated.

Table-VII.3: Foster et al (2007) Non-insertion/Semi-flexible

	Criteria	R	S	Comments	Median/ BM
1	Practitioner				2
2	Local points				4
3	Number of local points			Assumed	4-6 (3 or 7)
4	Meridian			One point from each of the 6 leg meridians	3
5	General health			SP-6, K-3	3
6	A-shi			trigger points	3
7	Relaxation			LI-4, LR-3	2
8	Other S&S				2
9	Unaffected knee			Not enough information	1
10	Back Shu points				1
11	At the elbow				1
12	Reduce needles			Some flexibility but limited	2
13	Degi				3
14	Depth of insertion				5-15mm (> 2)
15	Needle thickness				= 0.25mm (> 0.2)
16	Insertion time				20 mins (>15)
17	Session			Only six sessions	> 10 (≥ 6)
18	Follow-up sessions			None - 6 and 12 month data collection points	3
19	Moxibustion				2
20	Electro-acupuncture				2
21	Cupping				0
22	Total points			6-10	7-12 (5-16)

R - real acupuncture, S - sham acupuncture, BM - bench mark

Table-VII.4: Scharf et al (2006) Minimal/ Semi-flexible

	Mode	R	S	Comments	Median/ BM
1	Practitioner			Limited TCM training	x
2	Local points				4
3	Number of local points			7+	4-6 (3 or 7)
4	Meridian			Real no GB distal point. Sham on muscle meridians	3
5	General health			SP-6, K-3	3
6	A-shi				3
7	Relaxation			LR-3, LI-4	2
8	Other S&S			Only 16 possible points	2
9	Unaffected knee			If also affected opposite knee could be treated. Sham bilateral	1
10	Back Shu points			Real limited choice Bl-23	1
11	At the elbow			Real limited choice SI-8	1
12	Reduce needles			7 local points were obligatory	2
13	Deqi				3
14	Depth of insertion > 5mm			Shallow insertion is standard at acupuncture points close to the sham points	5-15mm (> 2)
15	Needle thickness				= 0.25mm (> 0.2)
16	Insertion time > 20 mins				20 mins (>15)
17	Session > 10			10-15	> 10 (≥6)
18	Follow-up sessions			Data collection 3 months after end of treatment	3
19	Moxibustion				2
20	Electro-acupuncture				2
21	Cupping				0
22	Total points			9-15	7-12 (5-16)

R - real acupuncture, S - sham acupuncture, BM - bench mark

Table-VII.5: Witt et al (2005) Minimal/ Semi-flexible

	Mode	R	S	Comments	Median/ BM
1	Practitioner				x
2	Local points			Limited TCM training	4
3	Number of local points			> 6	4-6 (3 or 7)
4	Meridian			sham 2 points on leg	3
5	General health				3
6	A-shi				3
7	Relaxation				2
8	Other S&S				2
9	Unaffected knee				1
10	Back Shu points				1
11	At the elbow				1
12	Reduce needles			Restricted ability to reduce number of needles	2
13	Deqi				3
14	Depth of insertion > 5mm				5-15mm (> 2)
15	Needle thickness				= 0.25mm (> 0.2)
16	Insertion time > 20 mins				20 mins (>15)
17	Session > 10				> 10 (≥ 6)
18	Follow-up sessions			A risk at weeks 26 and 52, not week 8	3
19	Moxibustion				2
20	Electro-acupuncture				2
21	Cupping				0
22	Total points			< 16	7-12 (5-16)

R - real acupuncture, S - sham acupuncture, BM - bench mark

Table-VII.6: Williamson et al (2007) No sham comparison/Semi-flexible

	Mode	R	S	Comments	Median/ BM
1	Practitioner				x
2	Local points				4
3	Number of local points			6-9	4-6 (3 or 7)
4	Meridian			Up to 3 points at acupuncturists discretion	3
5	General health			Up to 3 points at acupuncturists discretion	3
6	A-shi			Could use trigger points	3
7	Relaxation			LR-3	2
8	Other S&S			Western acupuncture so assumed none	2
9	Unaffected knee				1
10	Back Shu points			Western acupuncture so assumed none	1
11	At the elbow			Western acupuncture so assumed none	1
12	Reduce needles			Some flexibility in the protocol	2
13	Deqi			when possible	3
14	Depth of insertion			Not stated	5-15mm (> 2)
15	Needle thickness			0.3	= 0.25mm (> 0.2)
16	Insertion time			20.0	20 mins (>15)
17	Session			6 sessions	> 10 (≥ 6)
18	Follow-up sessions			Unclear as evaluation 6 weeks after treatment / post-operative	3
19	Moxibustion				2
20	Electro-acupuncture				2
21	Cupping				0
22	Total points			7-10	7-12 (5-16)

R - real acupuncture, S - sham acupuncture, BM - bench mark

Two protocols do not meet the primary benchmark of 10 sessions, but both meet the secondary benchmark of six (Foster et al., 2007, Williamson et al., 2007). Only one trial (Berman et al., 2004) tapered the sessions down to once a month. Follow-up treatments are an important part of practice, a median of three (Nō-18). This indicates that over half of TCM practitioners give follow-up treatments to 'more than 75% of their patients'. Whether the absence of follow-up sessions reflects of significant divergence from practice depends on when the outcome data is collected. Foster et al (2007) only provided six treatments, there was no follow-up, and data collection points were at six and 12 months. Hinman et al (2014) also collected data after 12 months. The length of time between last treatment and data collection appears to have compromised the external validity. In practice patients may be expected to return for a top-up treatment after roughly three months ([section-V.5.3](#)). Two other studies collected outcome data more than three months after the last treatment (Scharf et al., 2006, Witt et al., 2005). As there were no follow-up sessions external validity is compromised, even though the initial course was adequate ([tables-VII.4-5](#)). Vas et al (2004), Tukmachi et al (2004) and Mavrommatis et al (2012) conducted the evaluation shortly after treatment, so the lack of follow-up sessions is not relevant. Suarez-Almazor et al (2010) collected outcome data at three months.

The use of points to address the ben-root, or signs and symptoms (Nō-8), is difficult to incorporate within a semi-flexible protocol. The impact of its absence on external validity may also be related to the timing of data collection. Within the standard theory a ben-root pattern commonly associated with knee pain is Kidney Xu. This links osteoarthritis of the knee with the ageing process (Maciocia, 2005). A TBSA treatment may include not only certain acupuncture points but also lifestyle advice ([section-III.5.4](#)). It is possible that any potential benefit would only manifest in the long-term. Consequently, for short term outcome measures there may be minimal impact on the external validity even if Nō-8 is not included.

Table-VII.7: Vas et al (2004) Non-insertion/Fixed

	Criteria	R	S	Comments	Median/ BM
1	Practitioner			Acupuncturist	x
2	Local points				4
3	Number of local points			5	4-6 (3 or 7)
4	Meridian			ST40, SP6, KI3	3
5	General health			SP-6, KI-3	3
6	A-shi				3
7	Relaxation			LI-4	2
8	Other S&S				2
9	Unaffected knee			Not enough information	1
10	Back Shu points				1
11	At the elbow				1
12	Reduce needles				2
13	Deqi				3
14	Depth of insertion			Assumed	5-15mm (> 2)
15	Needle thickness			0.3	= 0.25mm (> 0.2)
16	Insertion time			Not enough information	20 mins (>15)
17	Session			12	> 10 (≥ 6)
18	Follow-up sessions			Evaluation only 1 week after treatment	3
19	Moxibustion				2
20	Electro-acupuncture			Some electric current, but needles were not inserted	2
21	Cupping				0
22	Total points			9	7-12 (5-16)

R - real acupuncture, S - sham acupuncture, BM - bench mark

Table-VII.8: Berman et al (2004) Non-insertion/Fixed

Criteria	R	S	Comments	Median/ BM
1 Practitioner			Acupuncturist	2
2 Local points				4
3 Number of local points			5	4-6 (3 or 7)
4 Meridian			BL-60, GB-39, KI-3, SP-6	3
5 General health			K-3, SP-6	3
6 A-shi				3
7 Relaxation			SP-6 has relaxation function	2
8 Other S&S				2
9 Unaffected knee			If also affected opposite knee could be treated.	1
10 Back Shu points				1
11 At the elbow				1
12 Reduce needles			Not flexible	2
13 Deqi			yes	3
14 Depth of insertion			0.3-1 inch	5-15mm (> 2)
15 Needle thickness			0.3	= 0.25mm (> 0.2)
16 Insertion time			20.0	20 mins (>15)
17 Session			23	> 10 (≥6)
18 Follow-up sessions			Tapered	3
19 Moxibustion				2
20 Electro-acupuncture				2
21 Cupping				0
22 Total points			9	7-12 (5-16)

R - real acupuncture, S - sham acupuncture, BM - bench mark

Table-VII.9: Suarez-Almazor et al (2010) Minimal/Fixed

	Mode	R	S	Comments	Median/ BM
1	Practitioner			Acupuncturists	x
2	Local points			Sham AC-LE-1 AC-LE-3 too close to local points	4
3	Number of local points			5	4-6 (3 or 7)
4	Meridian			Real only SP-6 possible	3
5	General health			Real only SP-6: Sham AC-LE-1 close to ST-36: AC-LE-3 close ST-34	3
6	A-shi				3
7	Relaxation			Real: SP-6 has relaxation function	2
8	Other S&S				2
9	Unaffected knee			Not clearly stated	1
10	Back Shu points				1
11	At the elbow				1
12	Reduce needles				2
13	Deqi				3
14	Depth of insertion			Can not be shallow and support electrodes	5-15mm (> 2)
15	Needle thickness			Real 0.22; Sham 0.2	= 0.25mm (> 0.2)
16	Insertion time			20	20 mins (>15)
17	Session			12	> 10 (≥ 6)
18	Follow-up sessions			6 weeks of treatment then evaluation at 3 months	3
19	Moxibustion				2
20	Electro-acupuncture			Stimulation of sham points	2
21	Cupping				0
22	Total points			6 plus ear point	7-12 (5-16)

R - real acupuncture, S - sham acupuncture, BM - bench mark

Table-VII.10: Tukmachi et al (2011) No sham comparison/Fixed

	Mode	R	S	Comments	Median/ BM
1	Practitioner				x
2	Local points				4
3	Number of local points			5	4-6 (3 or 7)
4	Meridian			Bl-57 , LR-3	3
5	General health				3
6	A-shi			None	3
7	Relaxation			LI-4 and LR-3	2
8	Other S&S				2
9	Unaffected knee			Not clearly stated	1
10	Back Shu points				1
11	At the elbow				1
12	Reduce needles				2
13	Deqi				3
14	Depth of insertion				5-15mm (> 2)
15	Needle thickness			0.3	= 0.25mm (> 0.2)
16	Insertion time			20-30	20 mins (>15)
17	Session			10	> 10 (≥6)
18	Follow-up sessions			Data collected at end of treatment so not applicable	3
19	Moxibustion				2
20	Electro-acupuncture				2
21	Cupping				0
22	Total points			7	7-12 (5-16)

R - real acupuncture, S - sham acupuncture, BM - bench mark

Table-VII.11: Hinman et al (2012) Sham laser/ Semi-flexible

	Mode	R	S	Comments	Median/ BM
1	Practitioner			Medical acupuncturist	x
2	Local points			not fixed	4
3	Number of local points			not fixed	4-6 (3 or 7)
4	Meridian			ST-40, BL-57/60, GB-35/36/41, SP-6, LR-3	3
5	General health			Limited choice	3
6	A-shi			in hamstrings	3
7	Relaxation			LR-3	2
8	Other S&S			unclear	2
9	Unaffected knee			-	1
10	Back Shu points			BL-21/22/23	1
11	At the elbow				1
12	Reduce needles			at beginning	2
13	Deqi			-	3
14	Depth of insertion			-	5-15mm (> 2)
15	Needle thickness			0.25	= 0.25mm (> 0.2)
16	Insertion time			20 minutes	20 mins (>15)
17	Session			8-12	> 10 (≥ 6)
18	Follow-up sessions			1 year follow up	3
19	Moxibustion				2
20	Electro-acupuncture				2
21	Cupping				0
22	Total points			not fixed	7-12 (5-16)

R - real acupuncture, S - sham acupuncture, BM - benchmark

Table-VII.12: Mavrommatis et al (2012) Fixed/ Non-insertion

	Mode	R	S	Comments	Median/ BM
1	Practitioner			Medical acupuncturist	x
2	Local points				4
3	Number of local points			6	4-6 (3 or 7)
4	Meridian			ST-40, KI-3, SP-6	3
5	General health			KI-3, SP-6	3
6	A-shi				3
7	Relaxation			LI-4	2
8	Other S&S				2
9	Unaffected knee			unclear	1
10	Back Shu points				1
11	At the elbow				1
12	Reduce needles				2
13	Deqi				3
14	Depth of insertion			Not stated bu deep enough to hold electrodes	5-15mm (> 2)
15	Needle thickness			0.30	= 0.25mm (> 0.2)
16	Insertion time			20 minutes	20 mins (>15)
17	Session			16	> 10 (≥ 6)
18	Follow-up sessions			data collected at 4 weeks after last treatment	3
19	Moxibustion				2
20	Electro-acupuncture				2
21	Cupping				0
22	Total points			9	7-12 (5-16)

R - real acupuncture, S - sham acupuncture, BM - benchmark

VII.3.2. Assessment of the risk of bias from sham procedures

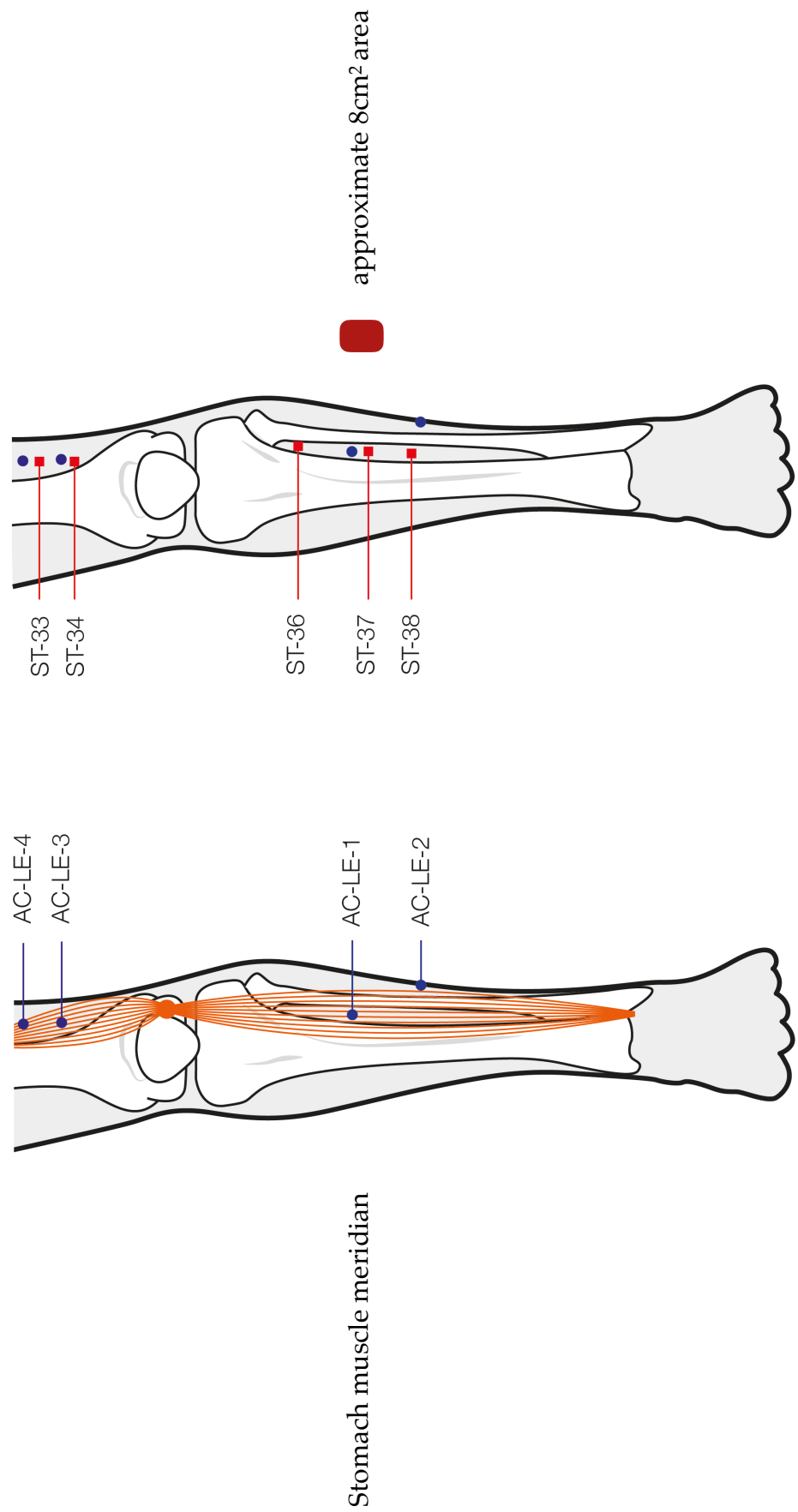
Eight trials used sham acupuncture procedures, four used non-insertion at acupuncture points (Berman et al., 2004, Foster et al., 2007, Mavrommatis et al., 2012, Vas et al., 2004), three used minimal acupuncture; i.e. non-acupuncture point shallow needling (Scharf et al., 2006, Suarez-Almazor et al., 2010, Witt et al., 2005), and one used a switched off laser (Hinman et al., 2014). Analysis of the non-insertion sham procedures is relatively straightforward because only a single component was tested, namely, skin penetration ([section-I.4.4](#)). If the acupuncture protocol lacks external validity due to the absence of a component, then this also absent in the control procedure. If the mode of point selection is absent, it is rated as red for the acupuncture and green for the sham procedure ([tables-VII.3/7/8](#)). Hinman et al (2014) has not been evaluated because it used a non-needle sham technique and the lack of clarity regarding the standardised list and 'other' points that could be used ([section-VII.2.2.a](#)).

Analysis of the minimal sham procedures is more complex. These procedures are characterised by shallow insertion at non-acupuncture points, and two particular problems have been identified (Appleyard et al., 2014). Firstly, TBSA include not only the standard meridians that many people are familiar with, but also muscle meridians. These follow approximately the same pathways as the standard meridians but are much wider. The muscle meridians are used to treat painful conditions (Maciocia, 2006). In addition, to stimulate these meridians shallow needle insertion is recommended. Secondly, acupuncturists do not locate points in exactly the same place, there is a degree of variation (Aird et al., 2002, Aird et al., 2000). If a sham point is too close to a real acupuncture point then in reality the two may overlap. Therefore, to assess the risk of bias it is necessary to evaluate whether each sham point is (a) too close to a relevant acupuncture point, and (b) situated within a relevant muscle meridian. The local (Nō-2) and meridian (No-4) modes of point selection appear to be the most important ([section-VI.4.3](#)). So, sham points that potentially fall into these categories are deemed to present a high risk of bias.

Suarez-Almazor et al (2010)

This trial utilised several sham points for which the risk of bias appears to be especially high. [Figure-VII.1](#) shows how the sham points used by Suarez-Almazor et al (2010) lie within the Stomach muscle meridian. AC-LE-2 appears to be on the edge but would lie within the Gallbladder muscle meridian. The issue of muscle meridians is probably unavoidable, if sham points are selected on the affected limb. However, the sham points of this trial are particularly egregious because of the proximity to highly relevant real acupuncture points. An elliptical area in the region of 8 cm² is required to ensure 95% confidence that the practitioners have inserted the needle in the 'same' location (Aird et al., 2002). A rectangle of approximately 8 cm² has been included to illustrate that the 'sham' points probably overlapped relevant acupuncture points. The Suarez-Almazor et al (2010) sham procedure has also been flagged red for stimulation and depth of insertion ([table-VII.9](#)). This is because they used an electro-acupuncture machine and briefly passed a current through the sham-points. The needles must have been inserted relatively deeply in order to hold the weight of the electrodes.

Figure-VII.1: Potentially active sham points



Scharf et al (2006)

The sham points on the affected limb used by Scharf et al (2006) were as follows:

One point between the gall bladder and stomach meridian on the distal part of the fibula, 2 cun above the Malleolus lateralis, in the direction towards the knee.

(Streitberger et al., 2004)

This sham point is certainly on the Gallbladder muscle meridian and probably too close to GB-39, which is indicated for knee pain (Ellis et al., 1991). GB-39 was used by Berman et al (2004) as a real point:

At each case, one point 2 cun and 6 cun above the malleolus medialis in the centre of the tibia surface area intracutaneous without periost contact

(Streitberger et al., 2004)

Both of these points are on the the Liver muscle meridian which runs along the tibia (Cheng, 1999, Ellis et al., 1991). The 6 cun point sits between LR5 ligou 蠡溝 (5 cun) and LR-7 中都 (7 cun) (Cheng, 1999, Zhao and Wang, 2007). These points are indicated for pain in the lower leg and Damp Bi (Ellis et al., 1991). Neither of these points are typically needled deeply (Cheng, 1999, Ellis et al., 1991). Therefore, the notion that shallow insertion at these points is a sham technique is incorrect.

One point in the middle of the thigh on a line from the centre of the patella to the spina iliaca anterior superior

(Streitberger et al., 2004)

This point lies on the Stomach muscle meridian (Cheng, 1999, Ellis et al., 1991).

Witt et al (2005)

The sham points on the affected limb used by Witt et al (2005) were as follows:

2 cun laterally (fibulary) and 1 cun caudal of BL-56

Six cun above the upper edge of the patella (between the Spleen and Stomach meridians)

(Brinkhaus et al., 2003)

The first of these points is situated on the Bladder muscle meridian, the other lies on the Stomach muscle meridian.

The Vickers et al studies undertook reasonable processes to develop the acupuncture and sham procedures. This begs the question as to why potentially active points were chosen. One potential reason is the diversity of TBSA and the ambiguity of ancient texts.

Clinical research requires standardisation, but this does not fit easily with the diverse practice of TBSA. Acupuncture points traditionally did not have the definitive locations that appear in modern textbooks (Appleyard et al., 2014). GB39 xuanzhong 悬钟 - is a good example of this problem. Some descriptions locate it 3 cun above the tip of the external malleolus, in the depression between the *posterior* border of the fibula and peroneus longus (Ellis et al., 1991). However, other textbooks put GB-39 *anterior* to the bone (Zhao and Wang, 2007, Zhao, 2002). In classical descriptions the acupuncturist should search for a pulse, 3 cun above external malleolus (Yang, 1601). Originally GB39 was described as only 1 cun above the malleolus (Huangfu, 282). To further illustrate the diversity this point is often referred to by another name *juegu* 绝骨, although Maciocia believes these are two different points: with *juegu* usually being higher up the leg (Maciocia, 2005)

Some textbooks describe the LR-5 and LR-6 as located on the midline of the medial surface of the tibia (Cheng, 1999, Zhao and Wang, 2007), in line with the Scharf et al (2006) sham point, others locate them on the medial edge of the tibia (Ellis et al., 1991). Locating the Liver meridian medial to the tibia, allows the needles to be inserted into a relatively fleshy area. If the point is located on the surface of the tibia, then insertion must be shallow. The Great Compendium simply describes LR-5 as being five cun above the medial malleolus, LR-6 as 7 cun above the medial malleolus in the middle of the tibia (Yang, 1601).

Potentially, the designers of the study thought that the Liver meridian was medial to the tibia. It is also possible that elements of TBSA have been overlooked because a placebo control was perceived as essential; in particular, the muscle and 'luo' meridians, which cover the body (Cheng, 1999, Ellis et al., 1991, Maciocia, 2006). The notion that TBSA posits needles must be inserted in definitive anatomical locations to be effective is fundamentally erroneous (Appleyard et al., 2014).

VII.3.3. Summary - external validity and risk of bias

Minimal acupuncture and non-insertion shams are different procedures and ought to be considered separately. The trials that use non-insertion procedures are relatively straightforward experiments to test needle insertion ([section-I.4.4](#)). In these trials, the central issue is to evaluate the external validity of the acupuncture. Among the trials that used non-insertion, the timing of outcome data collection and the number of sessions provided appear to be the most significant differences. Foster et al (2007) has the weakest external validity because not only were relatively few sessions provided, but also long-term outcome measures were used. The other studies (Berman et al., 2004, Vas et al., 2004) lack external validity with regards to the meridian mode of point selection (Nō-4).

The component analysis has shown that there is considerable heterogeneity among the minimal acupuncture procedures. The Suarez-Almazor et al (2010) sham procedure appears to have the highest risk of bias, Witt et al (2005) has the lowest and Scharf et al (2006) occupies the middle position. How well a protocol tests acupuncture versus 'placebo' depends on the limitations of the acupuncture and the risks of the sham. Again Witt et al (2005) appears to have the most external validity, Suarez-Almazor et al (2010) the least and Scharf et al (2006) in the middle. Suarez-Almazor et al (2010) should be excluded from future systematic reviews, because of the high-risk of bias of the sham procedure and the relative weakness of the acupuncture.

VII.3.4. Implications for future research

The analysis described above demonstrates that component analysis can be a useful tool in analysing the external validity of acupuncture protocols and the risk of bias from sham procedures. Component analysis should be used in the design of future research. This will enable greater clarity regarding the hypothesis that is being tested. Minimal acupuncture controls are particularly difficult to assess because of the ambiguity regarding whether the points are indeed sham. Clinical trials that use sham points should be able demonstrate that practitioners can consistently differentiate the sham from the nearest acupuncture point. Pre-trial tests could be carried out requiring practitioners to locate the points using marker pens that are invisible under normal lighting, but visible under ultraviolet light; similar to the research by Aird et al (2002). Trial reports should include discussion of the limitations of the fixed or semi-flexible acupuncture protocols. Researchers should analyse the risk of bias from the sham procedures. Perhaps more importantly, evaluations of both the experimental and control procedures should form part of subsequent systematic reviews.

VII.4. Generalisability of clinical research

The previous section assesses the external validity of fixed and semi-flexible protocols used in clinical research of osteoarthritis of the knee. Acupuncture practice is diverse ([section-I.3.3](#)). The extent to which these results can be generalised across the broad spectrum of practice is evaluated in this section. The systematic review ([chapter-IV](#)) examined the evidence-base for warm needle acupuncture versus needle acupuncture. The reporting and trial quality were found to be too low to draw any conclusions. Nonetheless, this section also considers the similarities and differences between acupuncture delivered in practice and clinical trials: both the high-quality western trials and the Chinese

research. The trials comparing warm needle to needle acupuncture, evaluated in [chapter-IV](#), are referred to as the Chinese studies.

VII.4.1. Generalisability assessed by modes of point selection

The practitioner survey found that there is a high degree of heterogeneity among acupuncturists who identified as using the same style. Nevertheless, there are broad differences between the three groups; medical acupuncture and the two TBSA groups, TCM and combined non-TCM. Medical acupuncture concentrates on local (Nō-2) and meridian (Nō-4) points ([section-VI.4.3](#)). The TBSA practitioners more frequently use points for General health (Nō-5) and to treat Other signs and symptoms (Nō-8). TCM differs from the combined group by placing greater emphasis on the local points (Nō-2) when compared to other signs and symptoms (Nō-8). All ten trials provided local (Nō-2) and some meridian (Nō-4) points that are the core treatment of many TCM and medical acupuncturists ([section-VI.4.3](#)).

Eight studies indicated that the acupuncture protocol was based on TCM (Berman et al., 2004, Brinkhaus et al., 2003, Foster et al., 2007, Mavrommatis et al., 2012, Streitberger et al., 2004, Suarez-Almazor et al., 2010, Tukmachi et al., 2004, Vas et al., 2004). Williamson et al (2007) report using medical acupuncture, yet, it is more reflective of TCM than Suarez-Almazor et al (2010) because of its capacity to freely select up to three points, whereas Suarez-Almazor et al (2010) only includes one meridian point (Nō-4). Hinman et al (2014) claim their protocol is a combination of TCM and medical acupuncture. Apart from two studies that incorporated some capacity to address the other signs and symptoms (Scharf et al., 2006, Witt et al., 2005), the others could easily be described as medical acupuncture rather than TCM.

VII.4.2. Generalisability assessed by needling parameters

Twenty four of the 32 Chinese studies describe the needle dimensions. The needles were typically 0.3 or 0.35mm in diameter, used by 14 and 8 studies respectively. A couple of studies used 0.25mm. The needles were usually 40 mm long, 50 mm were sometimes used, and one study used 75 mm (Tan and Feng, 2016) ([appendix-C.4](#)). Five of the Vickers et al trials used 0.3mm diameter needles (Foster et al., 2007, Mavrommatis et al., 2012, Scharf et al., 2006, Tukmachi et al., 2004, Vas et al., 2004), other trials reported using thinner needles, 0.25 or 0.22mm. The needle length was commonly 25-40mm. In general, the Chinese studies used thicker needles than the Vickers et al trials and typically used in UK practice.

Few Chinese studies report the actual depth of insertion; 25 mm for three studies and 1.5 cun for three studies - approximately the same depth. The Vickers et al trials tend to report shallower depths e.g. 5-25mm (Foster et al., 2007). Two Chinese studies used through needling (Yue, 2010, Zhang et al., 2009), this would have been particularly deep. Tan and Feng (2006) inserted the needles more deeply in the warm needle group, 2.5 cun as opposed to 1.5 cun. Needles were retained for 30 minutes in almost all Chinese studies. The Vickers et al trials generally left the needles for a shorter period time, 20-30 minutes. Chinese and Vickers et al trials both required deqi to be stimulated. Again acupuncture in the Chinese studies appears differ from U.K. practice and the Vickers et al trials: deeper insertion and longer retention time.

VII.4.3. Generalisability assessed by course of treatment

The mode number of treatments in the Chinese studies was 20, the range 9 to 40. For the Vickers et al trials the mean was 11.5 treatments, with only one study reaching the 20 treatments typical in Chinese studies (Berman et al., 2004). The Chinese studies also provided a higher intensity of treatment. Normally treatment was given every day with one or two days rest. The longest rest period was 'up to a week' (Yue, 2010, Zhang et al., 2009). Whereas the Vickers et

al trials offered one or two sessions per week, this was similar to the least frequent Chinese studies (Wang, 2007, Wang and Zhu, 2009).

VII.4.4. Comparison of acupuncture points used

Figure-VII.2 and figure-VII.3 show the frequency with which acupuncture points are cited for the local and distal points, respectively. Three sources are presented, the Chinese studies, the Vickers et al trials and the TCM textbooks (chapter-III). The charts show the percentage of citations within each set of sources: e.g. there were 9 textbooks, so if 3 cited a particular point, this would appear as 33%.

There is a strong emphasis on local points with all three sources citing the same local points multiple times (figure-VII.2). Particularly important are those points on the Stomach and Spleen meridians; that pass through the anterior aspect of the knee. The number of points actually used is not reported by any Chinese study, neither range nor mean. The number can be calculated from the protocols for a few studies, 3-4 (Li and Li, 2010, Yao et al., 2003) or 5-6 (Zhao et al., 2008). Probably a flexible approach was adopted in many of these studies. The number of local points used is reported by most of the Vickers et al trials, which typically used 5-6. There appears to be a high degree of similarity between the Chinese studies and the Vickers et al trials regarding the points and the number of needles used locally to the knee. These points are the same as those recommended by the TCM textbooks.

There is much greater heterogeneity in respect to the distal points. A total of 68 different distal points are cited. The TCM textbooks collectively cite 47 distal points, more than the Vickers et al trials or the Chinese studies, 37 and 12 respectively. Interestingly the Vickers et al trials cite the use of 17 distal points that are not recommended by any of the textbooks.

In the Chinese studies the most frequently cited point is Ren 4 *guanyuan*; six times. Twelve of the Chinese studies do not cite any distal points. The Vickers et al trials use more distal points. The Chinese studies do not report many

points that could be classified as meridian (Nō-4). Ironically this suggests Chinese acupuncture is more similar to medical acupuncture than TBSAs in the UK. It is possible the Chinese studies do not report all the points actually used. The character 等-*děng* which means etcetera is used in 20% of the Chinese papers. This may refer to local points only, it is often at the end of a list of local points, but whether this was the case is unclear. In one study 等-*děng* appears to indicate distal points as well (Xu, 2016).

VII.4.5. Summary

The acupuncture delivered within Chinese studies appears to be different from the Vickers et al trials. In particular, the number and frequency of treatments greater in the Chinese studies. The frequency of treatment in the Chinese studies was much higher than in UK practice, once a day compared to once a week. The needles were thicker 0.30 in the Chinese studies compared to 0.25; and retained for longer 30 minutes instead of 20. The reasons for these differences is unclear. They may be due to the limitations of time and expenses in the UK. Future research could investigate whether UK practitioners believe higher intensity of treatment would be more effective. It would also be of interest to investigate whether the thinner needles are thought to be more acceptable for western patients. The western studies report using more distal points, whether this is due to different reporting standards or actual practice requires further investigation. The use of points local to the knee is the common feature of both Western and Chinese protocols, as well as U.K. practice.

Figure-VII.2: Local points - citation frequency

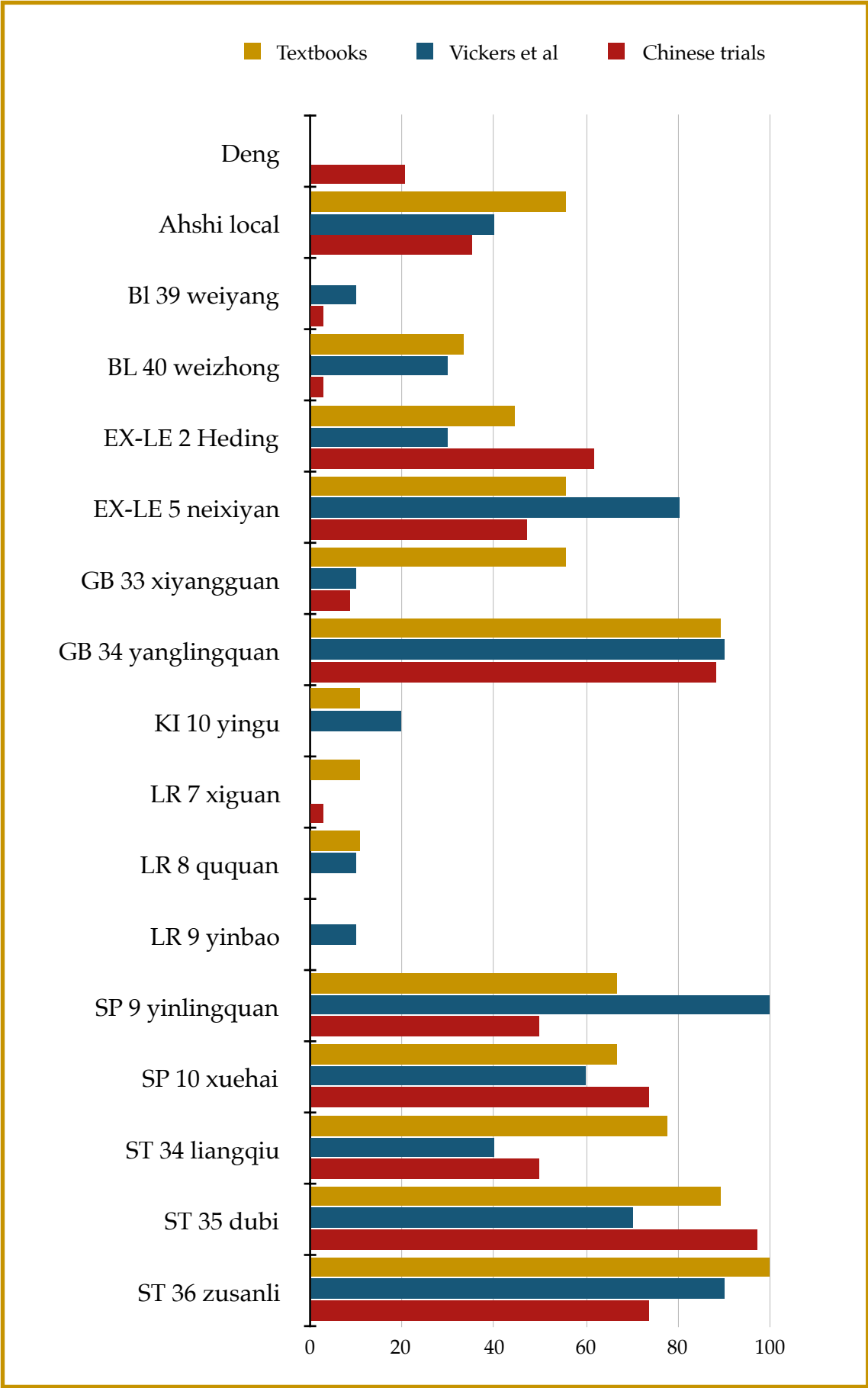


Figure-VII.3: Distal points - citation frequency

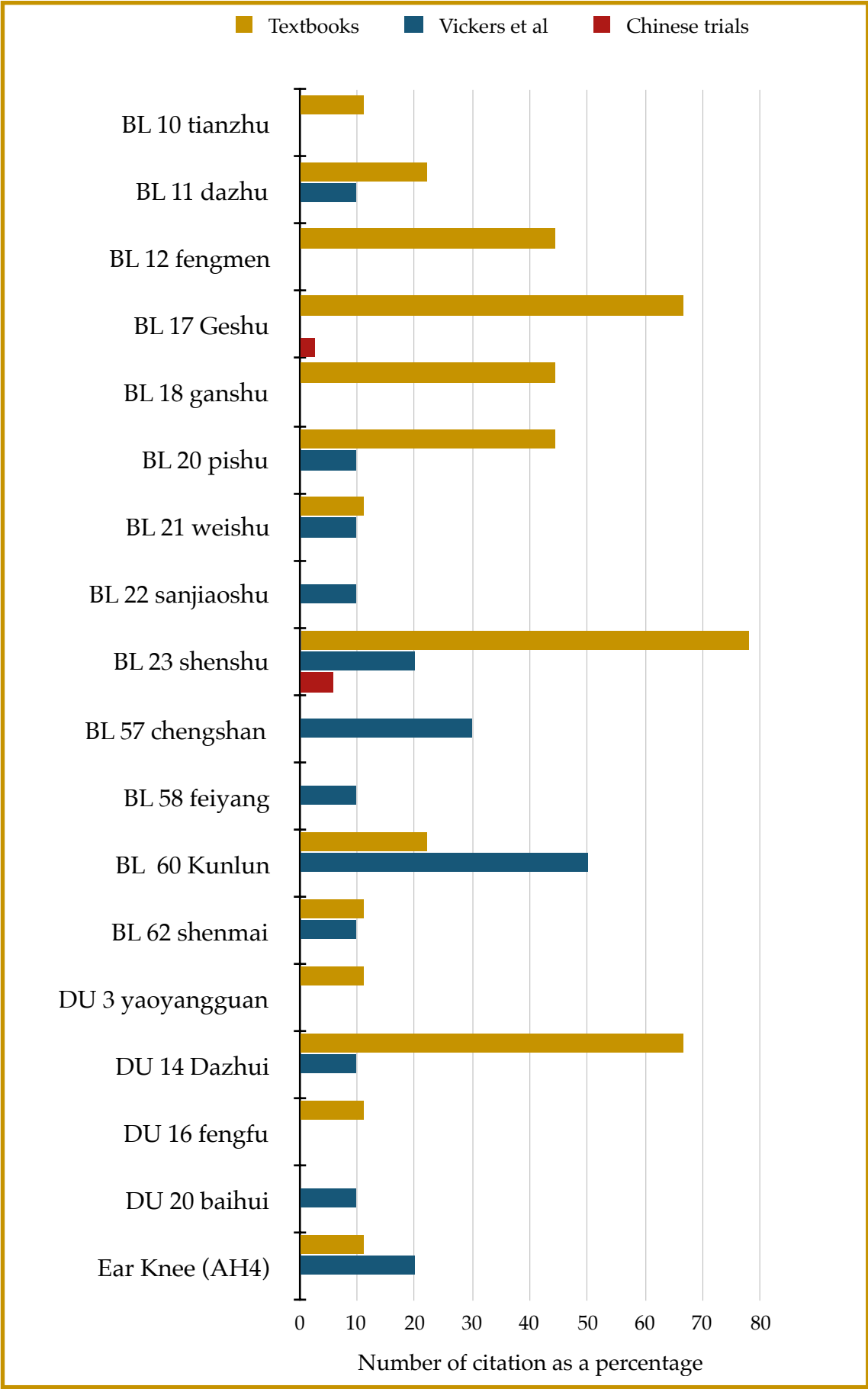


Figure-VII.3: Distal points - citation frequency (continued)

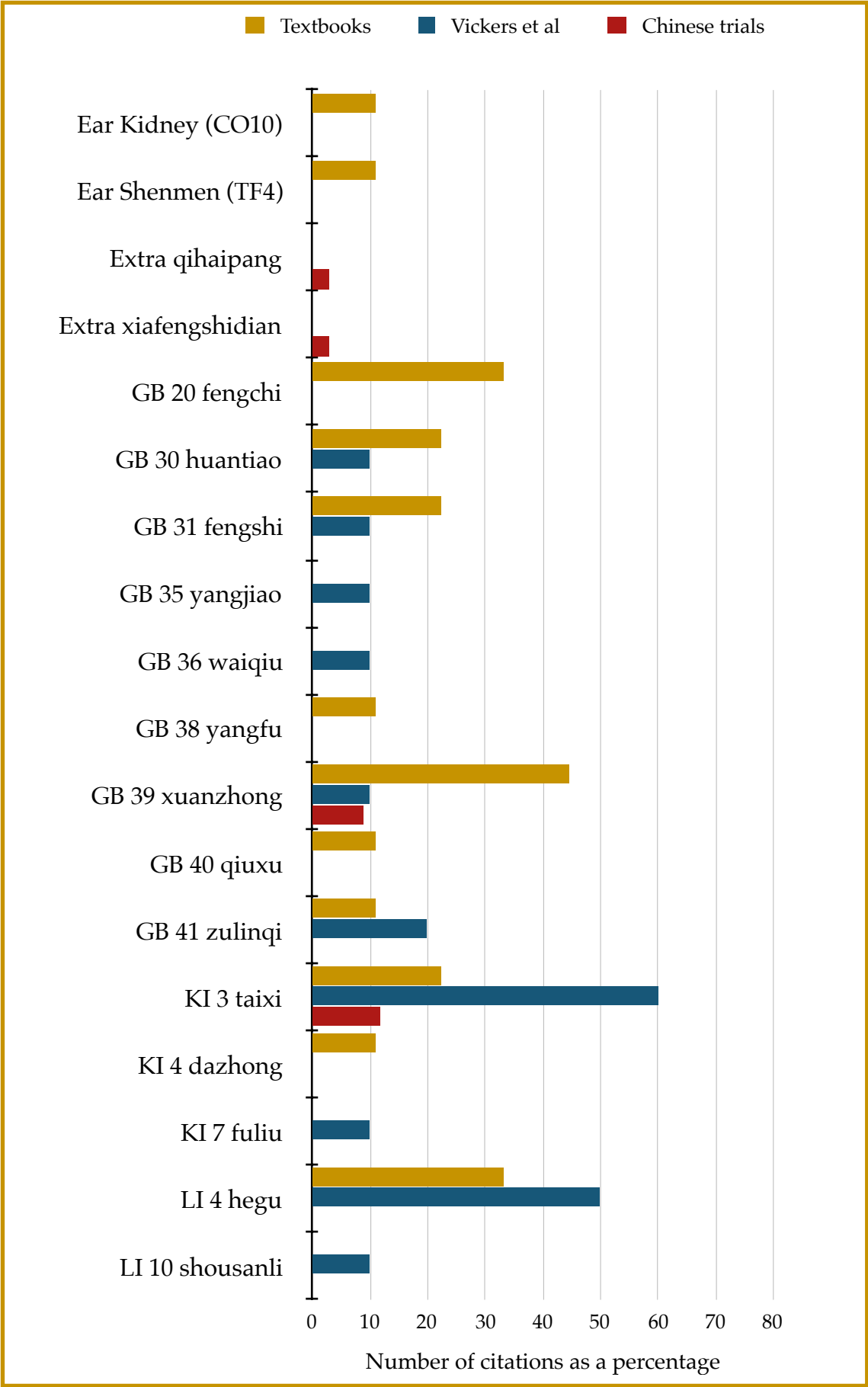


Figure-VII.3: Distal points - citation frequency (continued)

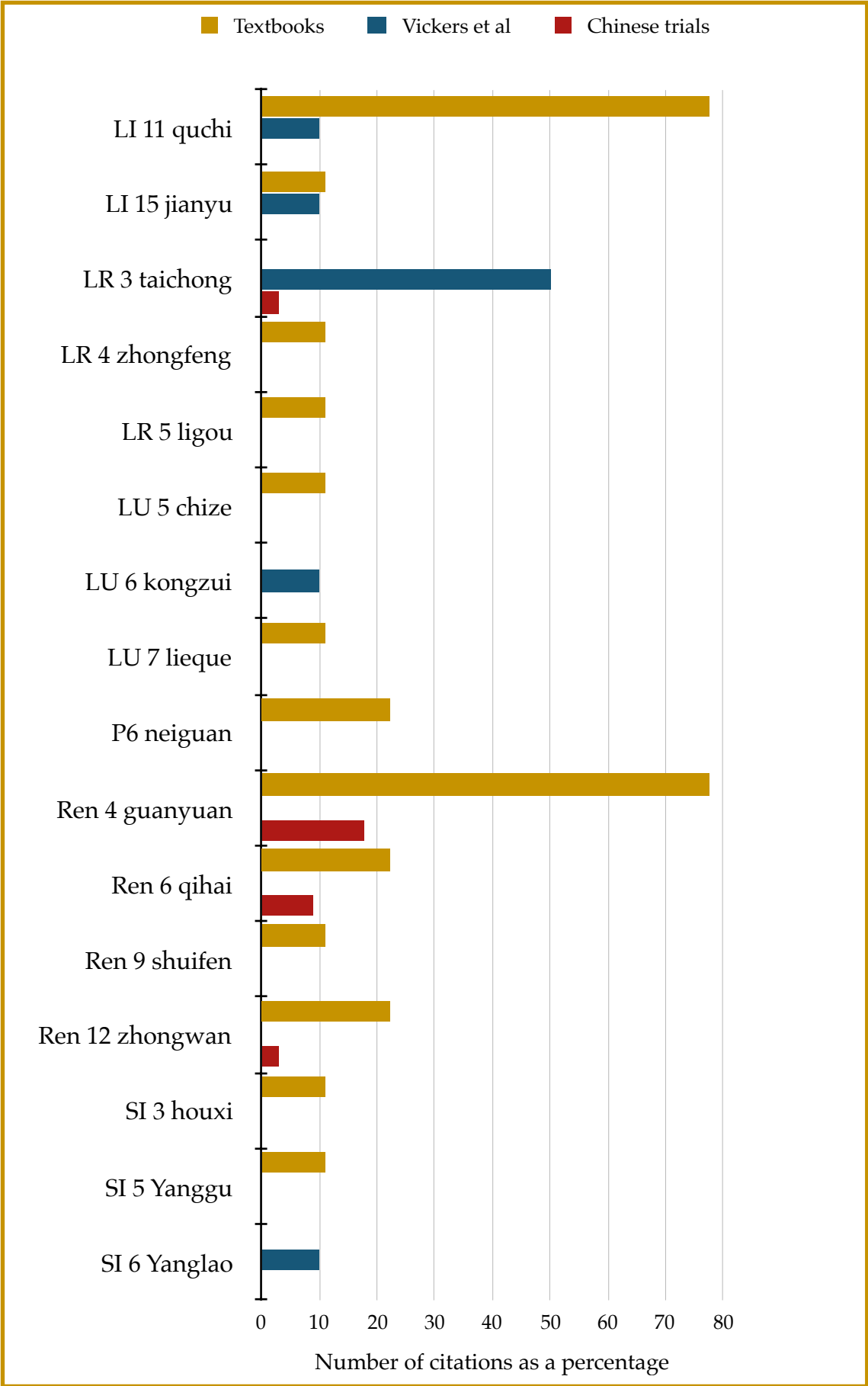
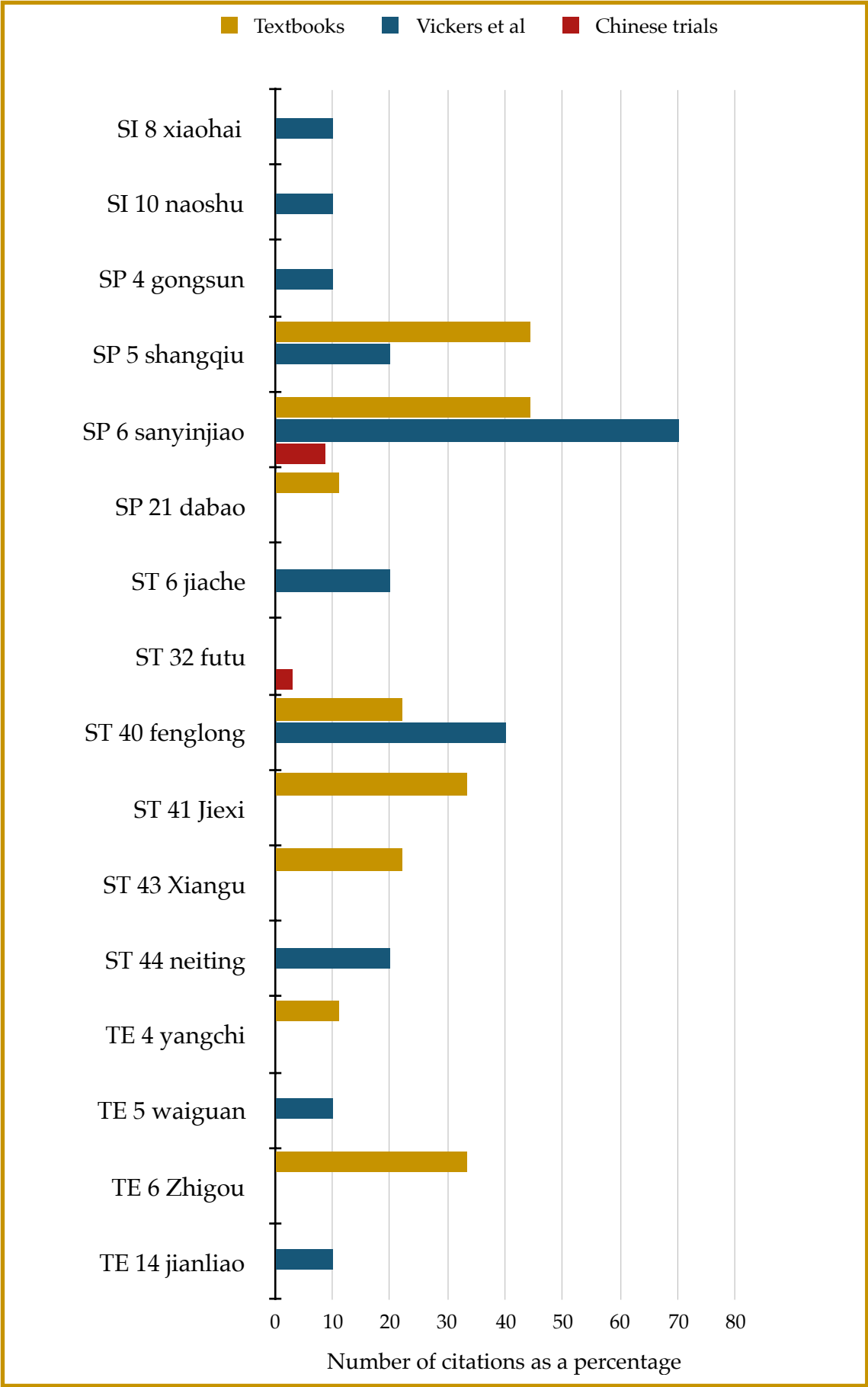


Figure-VII.3: Distal points - citation frequency (continued)



VII.5. Comparison of theory and practice

Chinese medicine has developed under a wide range of social and environmental influences ([section-I.3.3.b](#)). This raises the question as to whether comparing theory, practice and clinical research can shed any light on recent trends.

The literature review of the theory identified 47 acupuncture points, however, 24 of these were not used in any of the clinical trials ([figure-VII.3](#)). Notably, a set of points known as back-shu are indicated by almost half of the textbooks but seldom employed in clinical trials. Hinman et al (2014) include these points within their protocol but describe them as ‘segmental points’ (Hinman et al., 2012), suggesting that the underlying rationale was biomedical rather than TCM. The practitioner survey indicated that these points are not commonly used in practice ([section-VI.5.3.b](#)).

This begs the question as to why the back-shu points appear to be falling out of favour. Theoretically, bi-syndrome occurs if there has been an ‘external invasion’ ([section-III.5.1](#)), this implies exposure to wind, cold and damp. It seems reasonable to assume that these natural elements were more problematic historically in China than in modern Britain. Painful conditions may have had clear aetiologies related to exposure to the elements, but these have since diminished due to better clothing and building construction, office jobs rather than farming. Poor diet accompanied by lack of exercise leading to obesity, and an ageing population are commonly identified aetiologies of osteoarthritis of the knee (Bellamy et al., 2011, Silverwood et al., 2015). In the theory, back-shu points are not recommended specifically for the knee, but more generally for bi-syndrome regardless of the joint affected. The points are selected depending on whether the sub-pattern is considered to be Cold or Damp etc (Maciocia, 2005). Thus, they are closely linked to the notion of an external invasion. It is possible that practitioners no longer perceive the pain as being principally caused by environmental factors, therefore, tend not to use back-shu points. This change

in perception could be due to actual changes in the disease, or acupuncturists incorporating biomedical knowledge into their understanding of the disease.

Practitioners themselves may find it difficult to believe that inserting needles on the back will be effective in treating knee pain. They may also feel that the patient would not be satisfied with the treatment ([section-V.6.4.d](#)). More mundane reasons are also possible; it is simply inconvenient to insert local points on the anterior aspect of the knee and points on the back. This requires either two treatments, which is time-consuming, or the patient sitting/lying sideways, which can be uncomfortable. If the patient sits, it is not a very relaxing treatment. Further research is required to shed light on the clinical decision-making of acupuncturists which leads them not to select back-shu points as often as the theory might suggest.

VII.6. Discussion

This chapter has drawn together knowledge that was developed in the preceding phases of this project. New insights into the existing evidence-base have been generated. The analysis has demonstrated a new method of evaluating the external validity of fixed and semi-flexible acupuncture protocols, and the risk of bias from sham acupuncture procedures.

The acupuncture and placebo interventions of ten randomised controlled trials were assessed against 22 criteria. The evaluation indicates that the acupuncture protocols of some studies have greater external validity than others. The analysis has shown heterogeneity among the placebo acupuncture controls. In particular that the minimal acupuncture procedures are diverse, and some appear to exhibit greater risk of bias than others. Those conducting reviews in the future should consider whether trials such as Suarez-Almazor et al (2010), which appeared to have low external validity and a higher risk of bias from the sham procedure, should be excluded. Ideally, the evaluation of the

interventions would have been carried out by more than one researcher, who were blinded to the publications and results. In the future, it may be possible to begin to quantify some of the criteria ([table-VII.2](#)) in order to statistically evaluate their effects on outcomes.

The generalisability of clinical trials conducted both in China and the West has been evaluated. Generally, clinical research reflects medical acupuncture as much as, if not more than, TCM. Component analysis highlighted differences between acupuncture protocols that purported to be based on TCM. It was also not possible to clearly differentiate the medical acupuncture protocols from some TCM. This suggests that reporting the style of practice used in a clinical trials, as recommended by STRICTA (MacPherson et al., 2010), can only be of limited value. Component analysis provides a useful tool to compare protocols and evaluate whether or not the results should be combined.

The differences between theory and practice were also examined. This indicated that a group of acupuncture points known as back-shu appeared to have fallen out of use. Further research will be required to investigate the possible reasons for this apparent change in practice.

VIII. Pilot Study: Randomised Controlled Trial

试点研究

VIII.1. Introduction

The methodological framework of this project proposes that component efficacy trials should be utilised to advance acupuncture research ([chapter-II](#)). Moreover, knowledge of TBSA can help identify specific components to be tested. Moxibustion is an integral part of TBSA ([section-I.3.4](#)), yet, has not been used in methodologically rigorous trials (Vickers et al., 2017). As a result, moxibustion is a component of practice that requires further investigation. This chapter reports on a pilot study: the aim of which is to test a protocol to investigate the additional use of moxibustion ([box-VIII.1](#)).

Osteoarthritis of the knee was chosen as a suitable condition for the pilot study based on knowledge derived from phase 2 of this project. The inherent tensions between internal and external validity are examined in [chapter-II](#). A semi-flexible protocol is more likely to have external validity, if the bing-illness that is investigated has certain attributes ([box-II.2](#)). Moreover, a bing-illness where a semi-flexible protocol retains a high degree of external validity should be selected for efficacy trials of physical components ([section-II.4.5.b](#)). The narrative review of acupuncture theories suggests that osteoarthritis of the knee (knee pain) has the necessary attributes to test moxibustion ([section-III.5.6](#)). The condition selected should also meet other criteria ([section-II.5](#)). The disease burden and current treatment options should warrant the investigation of any therapy: evaluated in [section-III.2](#). Component efficacy trials should take place only after effectiveness has been established ([chapter-II](#)). The review of the evidence-base ([chapter-IV](#)) demonstrates that there is evidence of the effectiveness of needle acupuncture for osteoarthritis of the knee. A systematic review identified a number of clinical trials that had compared warm needle acupuncture to needle acupuncture, however, these trials have a high risk of bias. In addition, none of the trials used a method of blinding similar to the one developed for this pilot study ([chapter-IV](#)). Therefore there is a gap in the evidence-base.

There is no clear distinction between feasibility and pilot studies (Arain et al., 2010, Eldridge et al., 2016b). Pilot studies have been defined as small studies that test the components prior to conducting the main study. Whereas feasibility studies are used to estimate parameters needed for the design of the main study, such as the characteristics of proposed outcome measures and participant recruitment (Arain et al., 2010). This study used well-established outcome measures. Testing a specific design feature prior to conducting a full-scale RCT is a characteristic of a pilot study (Eldridge et al., 2016a). The protocol reported in this chapter uses a novel method for blinding participants and the acupuncturist. A key objective was to test these procedures, therefore the term - Pilot study - has been used.

The chapter describes the protocol design: randomisation and blinding procedures; recruitment and eligibility criteria; interventions. Clinical symptoms were measured using the Western Ontario and McMaster Universities Osteoarthritis Index, WOMAC[®] NRS 3.1 (Bellamy, 2016) and RAND-36 (version 1.0) (RAND, No date-a). The success of the blinding procedures is assessed using the Bang Index (Bang et al., 2010). Data on patient expectation, medication, safety, and TCM pattern differentiation were also collected. Participants were interviewed at the end of the course of treatment to gather information on their thoughts and feelings regarding various aspects of the study. The results and methods of analysis are presented. The strengths and limitations, along with ethical issues, of the study are debated. The inevitable compromises of a semi-flexible protocol with external validity are discussed within the context of the translational analysis presented in [chapter-VII](#). Recommendations for a prospective RCT are made.

The trial was registered at [ClinicalTrials.gov](https://clinicaltrials.gov) Ref: (NCT02680912) and the protocol published (Appleyard et al., 2016).

Box-VIII.1: Aim and objectives of the pilot study**Aim - to develop an RCT protocol to test the hypothesis:**

Warm needle acupuncture leads to greater clinical benefits for those with osteoarthritis of the knee than the use of needle acupuncture alone

Objectives

- ▶ Test the integrity of the study protocol. The study will enable the evaluation of the practicalities of the procedures and identify any problems that may arise from: implementing the inclusion/exclusion criteria; patient information and consent procedures; staff training; administration of outcome assessments; randomisation and allocation concealment.
- ▶ Assess the safety of warm needle acupuncture for OA of the knee
- ▶ Assess the acceptability of warm needle acupuncture among UK patients
- ▶ Collect qualitative data from participants and staff to support the development of the protocol for an adequately powered RCT
- ▶ Provide an initial indication of the effectiveness of warm needle acupuncture compared to needle acupuncture to inform a sample size calculation for an adequately powered RCT

VIII.2. Methods

VIII.2.1. Study Design

The study was a randomised, controlled, parallel-group 2-armed clinical trial. It was designed so that the outcome administrator, participants and primary acupuncturist would be blinded to patient allocation. Participants were randomised into two groups, they received either warm needle acupuncture or needle acupuncture.

The intention was to recruit 30 participants. Recommendations for sample sizes of pilot studies vary but the median number is 30 (Billingham et al., 2013). Smaller numbers can provide useful information regarding the acceptability of the procedures and highlight issues with implementation (Hertzog, 2008). It was hoped that 30 would provide an indication of the effectiveness of warm needle acupuncture to inform a sample size calculation for a subsequent full-scale trial. A number greater than 30 would have been impractical in terms of the available resources.

VIII.2.1.a. Randomisation

Participants were randomised into one of two groups, in 1:1 ratio. As this was a pilot study, with low numbers, participants were randomised in blocks of 10 to ensure periodic balance between the two groups. A statistician based at LSBU, one of the PhD supervisors, generated the randomisation sequence using a computer randomisation package. This sequence was expressed in terms of A or B. A different PhD supervisor chose whether group A or B would be warm needle acupuncture group, and prepared sealed envelopes ([figure-VIII.1](#)).

Figure-VIII.1: Group allocation process -flowchart

Random block sequence generated
by PhD supervisor

Trial Entrant Number	Block
1	A
2	B
3	B
-	-
-	-
-	-
10	B

Cards
Each participant may receive up to 12 treatments. 15 participants in each cohort.

The following were prepared:
15 sets of 12 cards labelled "Needle"
15 sets of 12 cards labelled "Warm"

Envelopes
12 envelopes were prepared for each trial entrant (they received up to 12 treatments)

The envelopes were labelled:
Trial Entrant Number (1,2,... 30)
Treatment Number (1,2,...12)

Different PhD supervisor:

- Decided whether Block A or B would be the warm needle acupuncture cohort.
- Inserted cards labelled with the corresponding group allocation into the 12 envelopes prepared for each Trial Entrant. Then sealed the envelopes.

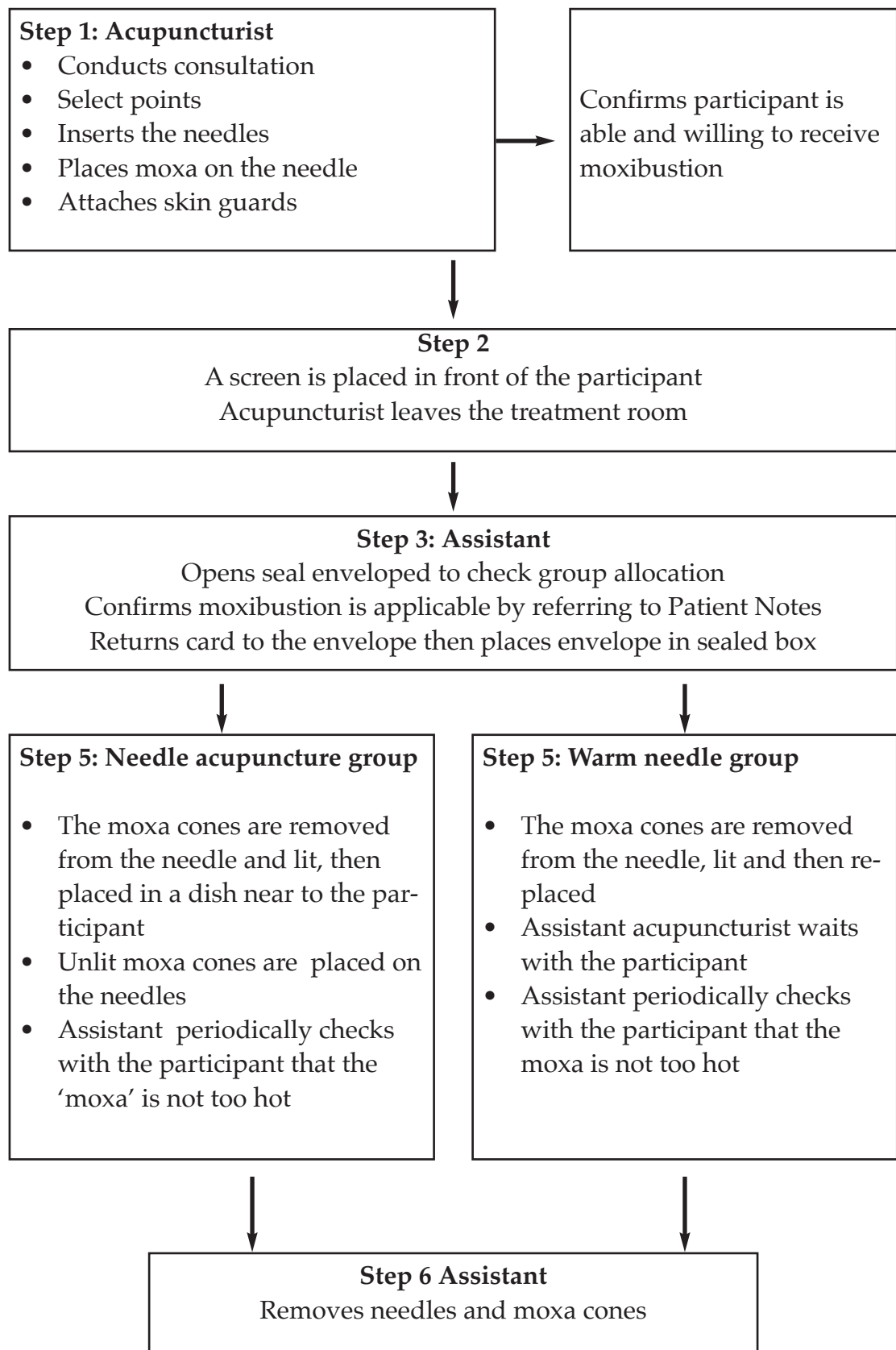
The sealed envelopes containing group allocation are stored with the preprepared Patient Notes Form

VIII.2.1.b. Blinding

A key innovation of this study was the blinding of both patient and practitioner. The researcher and those who administered the data collection were also blinded to group allocation. The blinding process required two practitioners; the acupuncturist and an assistant. The acupuncturist undertook the consultation, selected the points and inserted the needles. Blinding of the acupuncturist was important to ensure that no bias was created through additional care and attention to those in the warm needle group. The assistant lit the moxa and was the only person who was aware of group allocation.

A sealed envelope containing group allocation was opened at each treatment. This ensured that there was no information within the patient notes which identified a participant's group allocation that the acupuncturist could inadvertently discover. The process of preparing the sealed envelopes is illustrated in [figure-VIII.1](#).

The only difference in the procedures was that lit cones were placed on the needles of the treatment group whilst unlit moxa cones were placed on the needles of the control group. All patients saw the needles being inserted and the moxa cones being placed on the needles. Skin guards were placed at the base of the needle to reduce the immediate sense of heat on the surface of the skin ([figure-VIII.3](#)). The skin guards were specifically designed for this study and are made of a heat resistant and washable material. Without the skin guards blinding could not be achieved, as the radiated heat would be unmistakable.

Figure-VIII.2: Blinding procedure - flowchart

The acupuncturist carried out a consultation at each session as per normal practice. Participants were treated in an upright sitting position. Once the needles, skin guards and moxa were in place, a small table and screen were placed in front of the patient to prevent them from seeing their knees ([figure-VIII.4](#)). The acupuncturist left the room. The assistant then opened the envelope containing group allocation.

In the treatment group the assistant acupuncturist then removed one cone at a time and lit it. In the control group this action was mimicked to try and ensure the participants experienced the same sensations, i.e. all patients would hear the moxa being 'lit' and feel the cones being removed and replaced ([figure-VIII.5](#)). Although smokeless moxa was used patients would still be able to detect a faint smell. Therefore, in the control group moxa was lit and allowed to burn in a small dish, out of sight, at the participants feet ([figures-VIII.3/4](#)). At the end of the session all cones and ash were placed in a metal container with a lid, so the participants would not gain any visual clues. The needles and skin guards were then removed. Finally, the screen was taken away.

Figure-VIII.3: Close-up of warm needle acupuncture with skin guards



Close-up of needles with moxa and skin guards



Close-up of moxa in a dish. This moxa was lit during the treatment of participants in the needle acupuncture group, so that they would smell the moxa burning.

Figure-VIII.4: Treatment position with and without screen in place**Step 1:**

Needles are inserted, skin guard attached and moxa are placed on the needles.

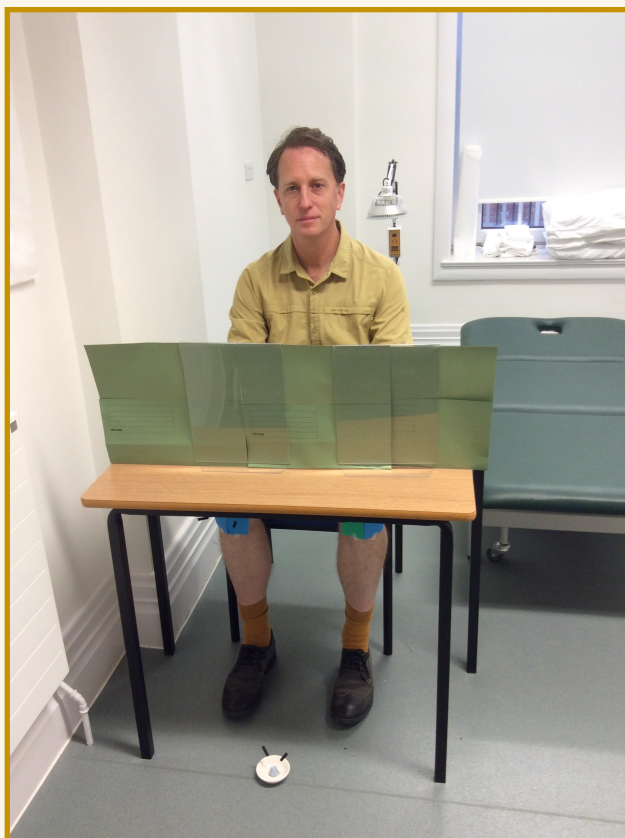
**Step 2:**

Table and screen put in position. Moxa placed at the feet of the participant.

Figure-VIII.5: The assistant lighting the moxa

The assistant removed each moxa cylinder in turn, lit it, and then replaced it on the needle.

If the participant was in the needle acupuncture group the moxa was not lit, but then it was removed and replaced so that the sensation would be the same.



VIII.2.2.Participants

VIII.2.2.a. Location of study

The study was conducted at the Confucius Institute of Traditional Chinese Medicine acupuncture teaching clinic based at London South Bank University. The trial lasted for 14 months, from April 2016 to May 2017.

VIII.2.2.b. Eligibility Criteria

The inclusion criteria incorporated the American College of Rheumatology (ACR) clinical criteria for diagnosing idiopathic osteoarthritis of the knee (American College of Rheumatology, 1986). Other elements of the inclusion criteria were designed to be broadly in line with previous studies of acupuncture for osteoarthritis of the knee (Streitberger et al., 2004, Witt et al., 2006). The eligibility criteria are set out in boxes [VIII-1](#) and [VIII-2](#).

Box-VIII.2: Inclusion criteria

Chronic pain in at least one knee joint during the last six months

At baseline the WOMAC[®] NRS 3.1 pain score must be ≥ 3 points (on a scale of 0–10)

The ACR criteria: in addition to the knee pain at least 3 of the following 6 must be present:

- Age > 50 years
- Stiffness < 30 minutes
- Crepitus
- Bony Tenderness
- Bony enlargement
- No palpable warmth

Ability to speak English

Signed consent form

Box-VIII.3: Exclusion criteria

A systemic disease of the musculoskeletal system

Bone tumour, bone tumour like lesions or metastasis

Bone fracture in the lower extremities during the last three months

Acute infection or osteonecrosis in the knee joint

Recent sprain injury to the knee joint

Surgery of the afflicted extremity during the last six months or planned surgery

Ongoing cortico-steroid therapy or cortisone injections in the past six weeks

Taking anti-coagulant medication

Coagulopathy

Other pain conditions which compels the patient to take analgesics for more than three days during the last four weeks

Addiction to analgesics, opiate or other drugs

Acupuncture treatment in the past 3 months

Dermatological disease within the acupuncture area impairing acupuncture treatment

Pregnant or breast-feeding patients

Inability to follow instructions or understand the consent form (insufficient command of language, dementia)

Participation in another clinical study

Ongoing legal proceedings concerning degree of disability

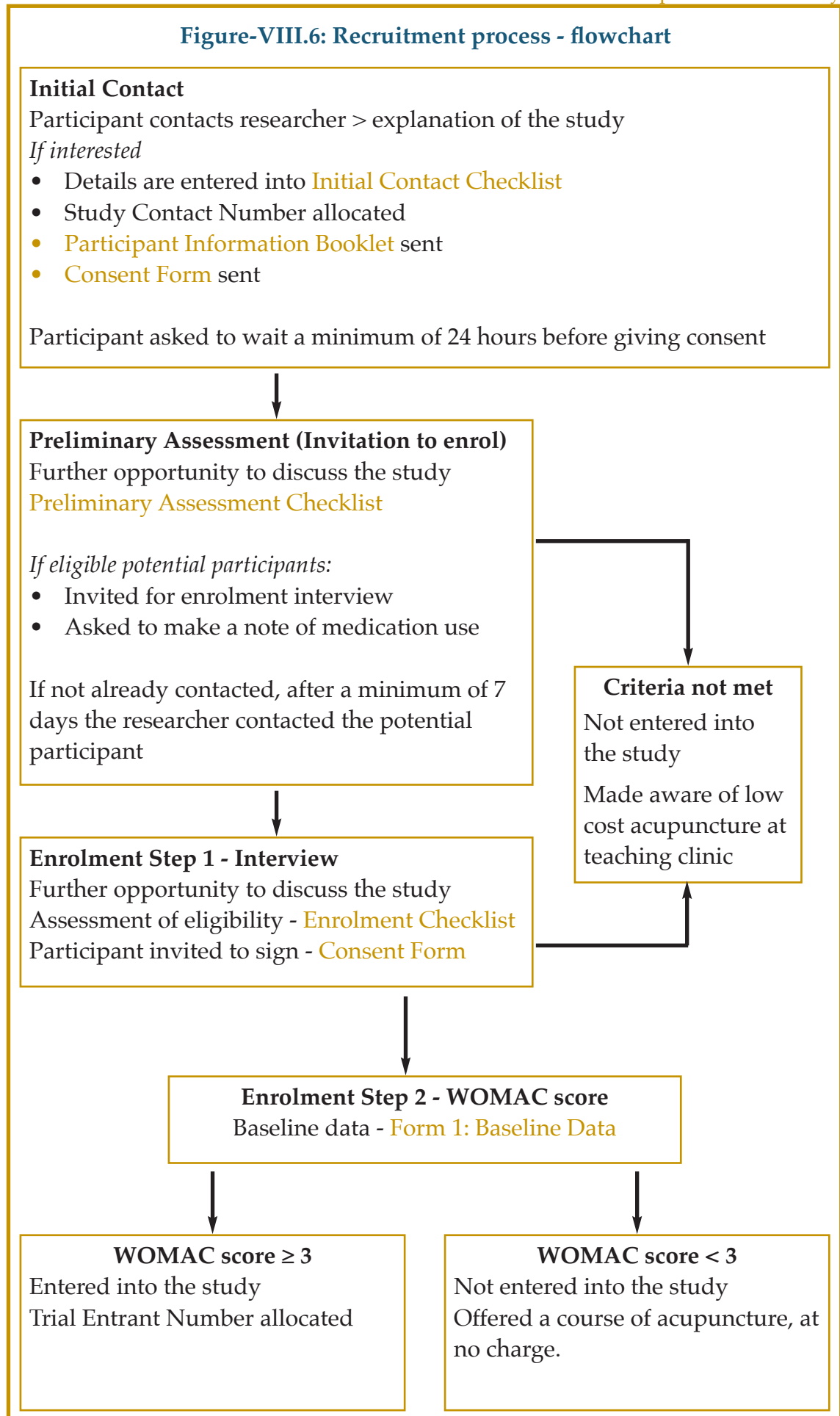
The following are excluded on the grounds that they would typically lead to a treatment different to the intervention protocols

TCM pattern differentiation of Heat Bi (not suitable for moxibustion)

VIII.2.2.c. Recruitment

Participants were recruited from the general public. LSBU's internal communication systems were used to publicise the study. Age UK-London and the U3A disseminated information about the study to their members. The publicity directed interested people to contact the researcher. The researcher would then explain the study and, if appropriate, send the participant information booklet ([appendix-F.10](#)) and consent form ([appendix-F.3](#)). Potential participants were asked to wait a minimum of 24 hours before contacting the researcher again, to ensure that they had had enough time to consider the nature of the study. If the potential participants had not responded after a week, they were contacted to enquire whether they wished to take part.

Once the potential participants had had the opportunity to consider the study and ask further questions, the researcher undertook a preliminary eligibility assessment, by phone. If suitable, individuals were invited to an enrolment interview. At the enrolment interview eligibility was rechecked ([appendix-F.2](#)), the consent form was signed ([appendix-F.3](#)), and the baseline WOMAC score calculated. Those who continued to meet all aspects of the eligibility criteria were enrolled into the study. Those who did not meet the criteria were offered a free course of acupuncture ([figure-VIII.6](#)).

Figure-VIII.6: Recruitment process - flowchart

VIII.2.3.Interventions

The only difference in the procedures was that lit cones were placed on the needles of the treatment group whilst unlit moxa cones were placed on the needles of the control group. [Box-VIII.4](#) briefly describes warm needle acupuncture for those unfamiliar with the procedure.

Box-VIII.4: Brief description of warm needle acupuncture

Moxibustion is the burning of moxa (*Mugwort*, *Artemisia vulgaris*) to warm specific parts of the body, including acupuncture points ([section-I.3.4](#)). Within the paradigm of traditional theories, one of the purposes of using moxibustion is to warm meridians and expel cold (Cheng, 1999). Osteoarthritis will typically be diagnosed as Cold Damp Bi syndrome in Chinese medicine. The use of moxibustion is indicated in the treatment of Cold Damp Bi syndrome ([chapter-III](#)). There are a number of ways to use moxibustion. Warm needle acupuncture (*wenzhen*; 温针) is where moxa cones are placed on the handle of the needle, after the needle has been inserted. Once lit, heat transmits along the shaft of the needle to the acupuncture point. In Chinese literature, acupuncture without the use of moxibustion on the needle is referred to as *danchun zhenci* 单纯针刺, which could be translated as simple needle insertion or *changgui zhenfa* 常规针法 regular or conventional needle method. In this thesis, the procedure has been referred to as ‘needle acupuncture’.

VIII.2.3.a. Intervention protocol - design process

The protocol has been published (Appleyard et al., 2016). The acupuncture interventions were designed by the researcher; an experienced TCM practitioner/ lecturer. They were based on: the literature review of the theory of acupuncture for knee pain ([section-III.5](#)) and the translational analysis of acupuncture for osteoarthritis of the knee ([chapter-VII](#)).

VIII.2.3.b. Acupuncture protocol

The same semi-flexible point selection protocol was used in both groups. Only points local to the knee were chosen. Points were selected according to the location of the pain. Between 4-6 points were used per knee, therefore 8-12 needles per treatment. Two points were used as the core treatment ST35 *dubi*, Ex-LE-5 *xiyan*. These points were omitted only if needling is not tolerated or inflammation/skin injury covered the acupuncture point. Other points could also be used from the following: Ahshi painful points local to the knee (*locus dolendi*), ST36 *zusanli*, GB34 *yanglingquan*, SP9 *yinglingquan*, ST34 *liangqiu*, SP10 *xuehai*, GB33 *yangxiguan*, LR7 *xiguan*, LR8 *ququan*, *hedong* Ex-LE-2.

Both knees were typically treated, even if only one knee was painful. This ensured participants received similar treatments and avoided the need to nominate a specific knee. It was thought possible participants may initially complain of pain in one knee, however, if the pain reduced they may start to focus on the other knee.

The practitioner survey indicated that *deqi* is not always obtained ([section-VI.4.4.b](#)) and stimulation is often adjusted according to patient sensitivity ([section-VII.4.7.a](#)). Therefore rather than mandating *deqi* be sought, the acupuncturists were allowed to stimulate the needles according to their usual practice. Practitioners did not manipulate needles inserted at ST35 *dubi*, Ex-LE-5 *xiyan*. This is inline with standard practice.

Ideally the knee should be at 90°, this facilitates needling of the key acupuncture points, ST35 *dubi* and Ex-LE 5 *xiyan*. Therefore, participants were treated in a sitting position ([figure-VIII.4](#)). Needles were retained for approximately 25-30 minutes. The needles were 0.25mm x 40mm, stainless steel single use Tewa 5JJ-Series (Straight Pipe handle).

Warm needle acupuncture

In addition to the acupuncture needles the warm needle acupuncture group received moxibustion. Smokeless moxa was used. Up to 4 points could be

selected to apply moxibustion to the needles per knee. Typically moxibustion was applied to ST 35 dubi, Ex-LE 5 xiyian and two other points. Two cones were sequentially applied to each needle. Smokeless Gu Gwan Moxa cones were used.

If the knee was red or there was palpable heat, then from a TCM perspective the pattern differentiation would be Heat Bi syndrome ([section-III.5](#)). Under these circumstances moxibustion would not typically be used. The protocol allowed for moxibustion to be omitted, if there were signs of heat or the participant felt particularly anxious, for example. The acupuncturist decided at each appointment whether moxibustion was applicable. The second cone of moxa could also be omitted, if it was deemed necessary by the acupuncturist or assistant. The decisions were recorded in the patient notes, at each treatment. If it was felt that the patient should not have warm needle acupuncture on 2 consecutive occasions then they would be withdrawn from the study. The participant could continue the course of acupuncture, if they wished. They were not withdrawn from the study, if only one cone per needle was administered.

Course of treatment

The practitioner survey indicated that treatments are typically given once a week, in the UK, although twice a week is not uncommon ([section-VI.5.6](#)). The Vickers et al (2017) trials conducted in the West offered treatment once or twice a week ([appendix-C.12](#)). The systematic review, however, showed that in China treatments were delivered more frequently, typically five times a week ([appendix-C.5](#)). The total number of treatments received on average (mean) in China was 17, in the Vickers et al (2012) trials the mean was 11.8. The practitioner survey suggested that in practice treatment may be extended over a long period of time, with patients returning for top-up treatments once every three months or so ([section-VI.5.6](#)). In this trial, participants could receive up to 12 treatments; 8 treatments in the first 4 weeks (twice a week), then 4 treatments in 4 weeks (once a week). This was identical to the pattern used in some of the

high quality trials that established the evidence base ([appendix-C.12](#)), and was similar to the initial intensity of practice in the UK ([section-VI.5.6](#)).

VIII.2.3.c. Practitioners

The acupuncture treatments were delivered by two practitioners. The researcher and a fellow member of staff at the Confucius Institute. Both were experienced practitioners with University degree qualifications in Chinese medicine acupuncture. Both were familiar with the TCM style of practice. The researcher was a member the British Acupuncture Council (BAcC), the other acupuncturist was a member of the Association of Traditional Chinese Medicine (ATCM). The assistants were also members of staff at the Confucius Institute/LSBU. All had undertaken degree level training in acupuncture in China or the U.K.

VIII.2.1.Data collection

Data were collected at baseline, midpoint (4 weeks), end of treatment (8 weeks) and follow up (16 weeks) ([figure-VIII.7](#)). In addition to the primary and secondary outcome measures, questions were also asked regarding patient expectation, medication, safety and the quality of blinding. At each treatment session the acupuncturist carried out a consultation as per usual practice. A preprepared form was used to record the clinical information ([appendix-F.5](#)).

Qualitative information was obtained from participants using semi-structured interviews. Information was gathered on the participants' experiences and perceptions of four areas; the trial process; needling sensations; the treatment process within the context of a trial; wider benefits or harms. The interviews were conducted at the end of treatment (week 8). The data collection forms are presented in the appendices:

- Form 1: Eligibility Baseline ([appendix-F.4](#))
- Form 2: Patient notes ([appendix-F.5](#))
- Form 3: Mid-point data - Summary check ([appendix-F.6](#))
- Form 4: End of treatment data - Summary check ([appendix-F.7](#))
- Form 5: Follow-up data - Summary check ([appendix-F.8](#))
- Interview schedule ([appendix-F.9](#))

VIII.2.1.a. Quantitative outcome measures

The primary outcome measure of the knee symptoms was the Western Ontario and McMaster Universities Osteoarthritis Index, WOMAC[®] NRS 3.1 (Bellamy, 2016). This is a validated outcome measure that has been extensively used in trials assessing acupuncture for OA of the knee ([appendix-C.13](#)). The secondary outcome measure was RAND-36 (version 1.0) (RAND, No date-a), which assesses the health related quality of life (HRQL).

VIII.2.1.b. Participant interviews

The interviews were semi-structured and sought to gather data on the participants' experiences and perceptions. Testing the methods and procedures has been described as a defining feature of a pilot study (Eldridge et al., 2016a, Thabane et al., 2010). The interviews were divided into four broad areas of interest.

The trial process: Open questions were asked regarding the recruitment process, information leaflet, consent forms and general running of the trial. The use of moxibustion is perhaps less well-known than needle acupuncture. Furthermore, the blinding procedure was quite elaborate. Therefore it was important to understand how well the information and recruitment process prepared participants for the study.

The treatment: The acupuncture provided within the context of a clinical trial may differ from that given in usual practice (Paterson and Britten, 2004). There is evidence to suggest that both patients and practitioners alter their behaviour in clinical trials (Paterson et al., 2008). This trial used a semi-flexible treatment protocol, which meant the points selected were not tailored to the individual as they might be in practice. For example no distal points were used. Questions were asked to ascertain how the participants perceived their treatment, and whether they felt it differed from what they would expect from normal practice.

The needling sensations: One of the key objectives of this study was to assess the blinding procedures. Participants were asked directly about their group

allocation. However, less direct more general questions regarding needle sensation were asked, as they might shed light upon any differences between the groups. The sensations, often described as deqi, that a patient feels at the acupuncture point may influence the therapeutic effect (Kim et al., 2008b, MacPherson et al., 2007). Acupuncture Sensation Questionnaires (ASQ) have been developed (Kim et al., 2008b). It was hoped information from this study could contribute to a better understanding of the sensation of having acupuncture.

Wider benefits or harms: The participants were asked to comment on how they felt the course of treatment had affected their lives, in order to try and capture information that may not have been covered elsewhere.

Figure-VIII.7: Data collection process - flowchart**Week 0 - Baseline**

- Baseline data: WOMAC / SF36 / Expectation / Medication
Baseline data (appendix-F.4)
- First Consultation: Patient notes (appendix-F.5)

Ideally the first consultation would be on the same day as the baseline data collection. If it was not possible to arrange the first treatment within 4 days then the baseline data would be retaken prior to the first consultation.

**Week 2**

Expectation Question (d) prior to the 4th treatment. This question was embedded within Patient notes (appendix-F.5)

**Week 4 - Mid-point**

WOMAC / SF36 / Allocation / Medication / Safety
Mid-point data (appendix-F.6)

**Week 8 - End of treatment**

WOMAC / SF36 / Allocation / Medication / Safety
End of treatment data (appendix-F.7)

Participant Interviews: Interview Schedule (appendix-F.9)

**Week 16 - Follow up**

WOMAC / SF36 / Medication / Safety
Follow-up data (appendix-F.8)

VIII.2.1.c. Patient expectation

Previous studies have suggested that expectation plays a role in identifying patients who respond to acupuncture (Linde et al., 2007). Four questions were asked to assess patient expectation (box-VIII.5). Three of these (a,b,d) had been used in previous studies (Linde et al., 2007). Question (c) used the same structure as the Linde et al (2007) questions but enquired about moxibustion. Questions (a,b,c) were asked at the initial consultation, question (d) asked at the fourth session.

Box-VIII.5: Expectation questions**a) How effective do you consider acupuncture in general?**

Answer options:

very effective / effective / slightly effective / not effective / don't know

b) What do you personally expect from the acupuncture treatment you will receive?

Answer options:

cure / clear improvement / slight improvement / no improvement / don't know

c) Do you think warm needle acupuncture will be more effective than needle acupuncture for your knees

Answer options:

much better / better / slightly better / the same / don't know



The following question was asked at the fourth session.

d) How confident do you feel that this treatment can alleviate your knee pain?

Answer options:

very confident / confident / slightly confident / not confident / don't know

VIII.2.1.d. The quality of blinding

Participants were asked the following question regarding group allocation at the midpoint and end of treatment. Those who answered warm needle or needle acupuncture were asked to state why.

Do you think you are receiving warm needle acupuncture or needle acupuncture?

Answer options

- a. Warm needle b. Needle acupuncture c. Don't know

VIII.2.1.e. Medication

Participants were asked to make a note of their use of medication ([appendix-F.2](#)). The amount of medication used was recorded at baseline-week 0, midpoint-week 4, end of treatment-week 8 and follow up-week 16 ([appendix-F.4/6-8](#)).

VIII.2.1.f. Safety Evaluation

Adverse events were monitored at each consultation, by the acupuncturist and recorded in the patient notes. Participants were also asked an open question at the midpoint and end of treatment:

Have you experienced any side effects or adverse events because of the acupuncture?

VIII.2.1.g. TCM pattern differentiation

The acupuncturists made a full TCM differential diagnosis at the first treatment session. This was recorded on the Patient notes ([appendix-F.5](#)). If sufficient data was collected, the relationship between differential patterns and the dependent variable would be analysed.

VIII.2.2. Analysis

The primary outcome measures was WOMAC® NRS 3.1 (Bellamy, 2016). The secondary outcome measure was RAND-36 (version 1.0) (RAND, No date-a). These were both considered interval variables, the change in score would be analysed using t-test; provided the distribution of changes were normally

distributed. Paired samples t-test was used to analyse changes over the four-time points. Independent samples t-test was used to investigate the relationship between the categorical explanatory variable and change in WOMAC and RAND-36 (v1.0). The assumption of normality was tested using the Shapiro-Wilk test, and homogeneity of variance using the Levene test (Field, 2009, Ghasemi and Zahediasl, 2012). The study only recruited 14 participants, therefore it was underpowered for detailed investigation of change over multiple time points or for investigation of possible explanatory variables.

The WOMAC® Osteoarthritis Index is a tri-dimensional, disease-specific, patient reported outcome measure (PRO) (Bellamy, 2015). The dimensions are pain, stiffness and function. The version used in this study was the NRS 3.1 which is an 11 point likert scale, 0 - 10, where 0 is no pain and 10 is extreme pain. The index consist of 24 questions, 5 pain, 2 stiffness and 17 physical function.

There are a number of ways in which three sub-scales can be combined into a single score (Bellamy, 2015). In this study the three sub-scales were first normalised to 0-10, these were then combined in 1:1:1 ratio. The resulting total score 0-30 was then normalised to 0-10 scale. In other words, three sub-scales have been presented as mean scores on a 0-10 scale, as has the total score.

The impact that participant expectations may have had on outcomes was analysed by converting the WOMAC pain scores and responses to the expectation questions into binomial data and using the Fisher exact test (Linde et al., 2007). Each of the expectation questions have five possible answers (box-VIII.5), these were converted to a binomial variable; high-expectation or low-expectation. The two most positive responses were considered high-expectation e.g. very effective/ effective. The remaining three responses were categorised as low expectation. The WOMAC pain scores were converted to percentage change from baseline and then to responder rates. The criterion for 'responder'

was a change from baseline of $\geq 50\%$; this was in line with previous studies (Linde et al., 2007, Vickers et al., 2012)

The RAND 36-Item Health Survey measures eight health concepts: physical functioning, bodily pain, role limitations due to physical health problems, role limitations due to personal or emotional problems, emotional well-being, social functioning, energy/fatigue, and general health perceptions. There are different scoring systems and versions of the survey, this study used RAND 36-Item Health Survey (Version 1.0) (RAND, no date-b). The scales of the eight health concepts can be combined to form two aggregate scales, mental health and physical health. There are different methods of combining the data (Laucis et al., 2015). The methods described as the physical health composite (PHC) and mental-health composite (MHC) have been used in this study. A higher score on any of the scales of RAND-36 indicates better health.

The interviews were carried out in the week after the last treatment. The information obtained from the interviews was thematically analysed using framework analysis (Gale et al., 2013, Ritchie and Spencer, 1995). This provided qualitative information regarding the thoughts, feelings and experiences of the study from the participant perspective. This information was used to evaluate the practicality of the procedures and identify any problems that may have arisen from: implementing the inclusion/exclusion criteria; patient information and consent procedures; staff training; administration of outcome assessments; randomisation and allocation concealment.

Blinding was also assessed using Bang's Blinding Index (BI) (Bang et al., 2010). The BI produces a number between 1 and -1, 1 indicates all guesses were correct and -1 all guessed incorrectly; random 50-50 guesses would be 0. The index is calculated separately for experimental and control group (Bang et al., 2010).

Statistical analysis was carried out using IBM® SPSS® Statistics Version 21. The RAND-36 scores were converted to the PHC and MHC aggregate scores using

the spreadsheet available at: http://journals.lww.com/jbjsjournal/Abstract/2015/10070/Scoring_the_SF_36_in_Orthopaedics___A_Brief_Guide.10.aspx

VIII.2.3.Ethics

Ethical approval was obtained through London South Bank University's Ethics Committee ([appendix-F.1](#)). The World Medical Association Declaration of Helsinki and Social Research Association: Ethical Guidelines (World-Medical-Association, 2013); and London South Bank University Research Ethics Committee Code of Practice for Research Involving Human Participants (LSBU, 2011) were used for guidance. The study was designed with reference to British Acupuncture Council Codes of Practice (British Acupuncture Council, 2016); information available at the websites of the Health Research Authority (Health Research Authority, 2014) and National Institute for Health and Research (National Institute for Health Research, 2014).

The LSBU code of practice stipulates that participants should have autonomy and have 'the right to think independently and act freely to decide to participate' (LSBU, 2011). The participants were provided with an information booklet; containing information about the objectives of the study, what was expected of them, how long their participation would last, any risks involved, how the data would be handled and who could be contacted for further information ([appendix-F.10](#)). Participants were given up to 1 week to consider whether they wished to participate in the study ([figure-VIII.6](#)).

The researcher spoke with the participants to provide information on the objectives, risks, use of the data and inconveniences of the study. It was clarified that participation was voluntary and that the participant would be free to withdraw from the study at any time, without reason giving a reason. If the participant agreed to take part in the study, the researcher obtained 2 signed consent forms ([appendix-F.3](#)). The participant kept one copy the other was retained by the researcher. Participants were also be asked to consent to have acupuncture immediately prior to the first treatment. This was done by signing

a consent form embedded within the patient notes; this is standard practice at the LSBU teaching clinic ([appendix-F.5](#)). No participant was recruited within 24 hours of initial contact, to ensure they had enough time to read and absorb all the information.

The LSBU code of practice stipulates that the research should have beneficence (LSBU, 2011). Participants received a course of acupuncture treatment free of charge, although no funds were available to cover transport costs and other out of pocket expenses. Needle acupuncture has been shown to offer clinically relevant benefits for patients with OA of the knee (Manheimer et al., 2010, Vickers et al., 2012). It was thought unlikely that warm needle acupuncture would be inferior to needle acupuncture.

Risks

Both procedures are part of the standard practice of acupuncture in the UK, China and around the world. Consequently, risks from the treatment were thought to be equivalent to receiving acupuncture from a qualified practitioner in everyday practice. Acupuncture has been shown to be a safe intervention (MacPherson and Thomas, 2005, MacPherson et al., 2001, White, 2004, Witt et al., 2009).

In an observational study of 229,230 patients who received an average of 10 treatments, 8.6% of patients reported at least one adverse event. The most common adverse event was bleeding or bruising 6.1% (58% of the total). Pain was reported by 1.7%, nausea 0.15% and dizziness by 0.22% (Witt et al., 2009). These adverse events are short term and do not typically require further treatment. The most serious event reported was pneumothorax. Pneumothorax could not occur in this study due to the location of the needles. The participants information sheet explained the risks of acupuncture ([appendix-F.10](#)).

All acupuncturists who took part in the study were trained to a minimum of degree level, and were members of either the British Acupuncture Council (BAcC) or Association of Traditional Chinese Medicine (ATCM) with current

first aid certificates. All practitioners followed the safety guidelines of the British Acupuncture Council Guide (British Acupuncture Council, 2016). Single-use, pre-sterilised, disposable, solid acupuncture needles were used. Sharps boxes were clearly labelled and conformed to BS 7320:1990.

Adverse events were monitored by the acupuncturists throughout the course of treatment and recorded on the patient notes ([appendix-F.5](#)). Adverse events were also recorded in the clinic logbook as required by the BAAC codes of practice (British Acupuncture Council, 2016). Assessment as to whether or not further medical attention was required would be made by the acupuncturists, in line with BAAC codes of practice (British Acupuncture Council, 2016). Adverse events were also discussed in the qualitative interviews ([section-VIII.2.1.b](#)).

There were no anticipated risks to any researchers. The acupuncturists faced the same risks (cross infection) as normal acupuncture practice. If any interviews were conducted at the participants home, the researcher would ensure that contact details of where he had gone and estimated time of return were left with other members of the research team. He would have taken a mobile phone and adhered to the university lone workers policy.

Peer Review

The protocol for this study was peer-reviewed by the British Acupuncture Council as part of the grant award process.

Data Storage

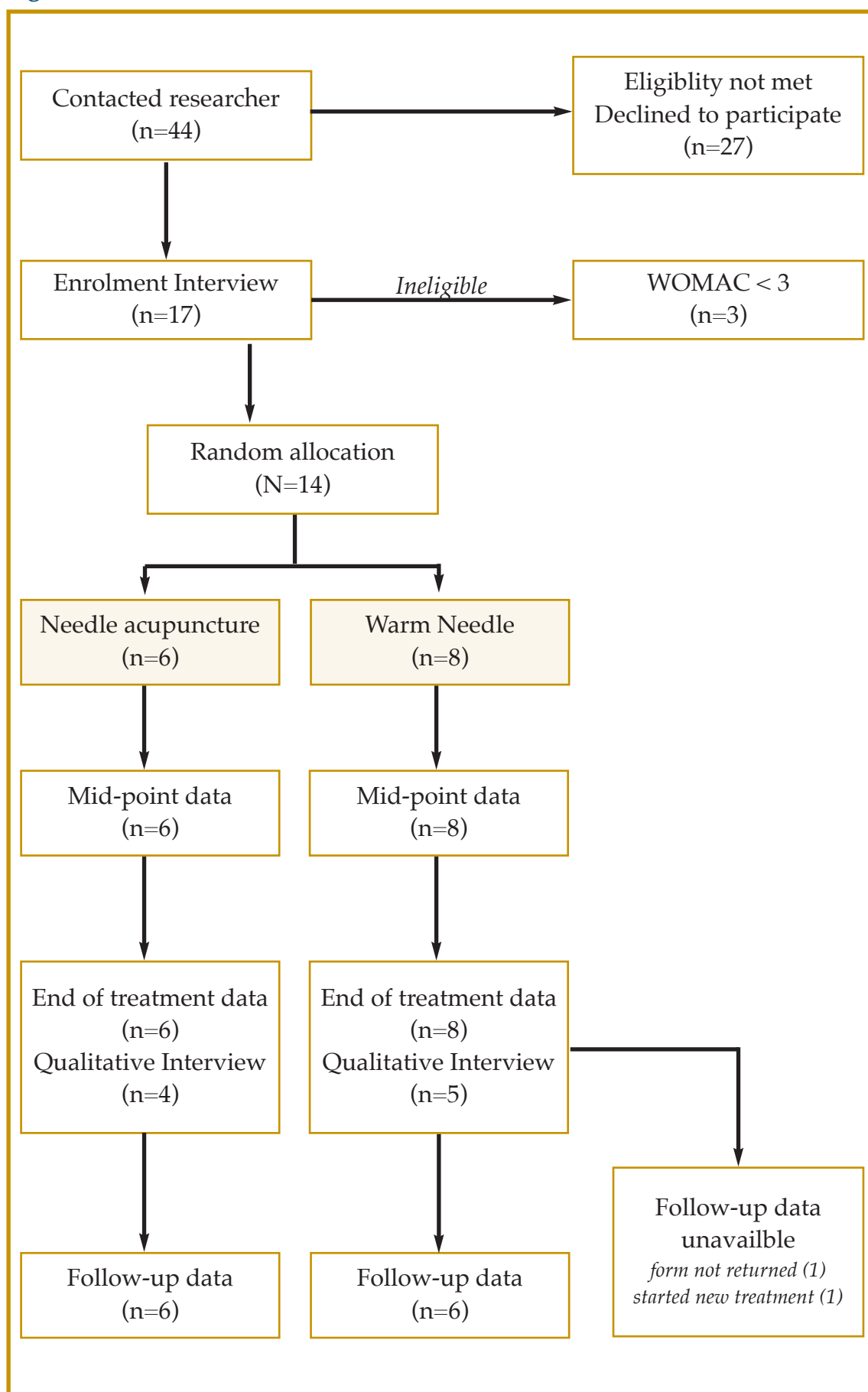
All information collected during the course of the research has, and will be, kept strictly confidential. The notes are kept in a locked filing cabinet within the university. The data will be stored on a password-protected computer at LSBU until 2023. The notes and data will then be deleted/destroyed.

VIII.3. Results

VIII.3.1. Recruitment, attendance and data collection

A total of 44 people contacted the researcher. Of these, 27 either did not meet the eligibility criteria (n=21) or declined to participate (n=6). Seventeen people were invited to an enrolment interview. Three were found to have a pain score of below 3 at baseline. Therefore, 14 people were enrolled into the study. Eight were randomly allocated to the warm needle group and six to the control group. Most participants in the warm needle group received all 12 treatments (7/8), the remaining participant attended 11 times. In the needle group three had all 12 treatments, two had 11, and one had 10.

Data were collected for all 14 participants, at mid-point (week 4) and end of treatment (week 8). No data were collected from two participants at the follow-up time point (week 16), both in the warm needle group ([figure-VIII.8](#)). One had an operation and so no longer met the eligibility criteria. The other failed to return completed questionnaires. Interviews were conducted with 9/14 participants: warm needle group n=5, controls n=4. In the majority of cases this was immediately after the last treatment 7/9. One interview was conducted 3 days after the last treatment. One participant was unable to attend the arranged interview time, but provided written answers to the prompts of the interview schedule ([appendix-F.9](#)).

Figure-VIII.8: Recruitment, enrolment and data collection - flowchart

VIII.3.2. Baseline characteristics

The majority of participants were female, 11 out of 14. Most had pain in both knees, 10/14. The average age was 54.4 (SD 12.0) years and the mean duration of the knee pain was 10.3 years (SD 8.1). There was a wide range in the duration of pain, particularly in the warm needle group 1-30 years. There were no statistically significant differences in the demographic data at baseline (table-VIII.1).

Table-VIII.1: Baseline demographic characteristics and pattern

Variable	Warm mean (SD)	Needle mean (SD)	Total
Participants n	8	6	14
Women n	6	5	11
Mean age (SD) years	56.9 (10.3)	51.0 (13.2)	54.4 (12.0)
Mean duration of pain (SD) years	9.6 (9.6)	11.2 (5.4)	10.3 (8.1)
Range duration of pain years	1-30	5-21	1-30
Both knees affected n	6	4	10
Bi-syndrome pattern			
Cold Bi	0	2	2
Damp Bi	4	2	6
Cold Damp Bi	2	2	4
Qi and Blood Stagnation	2	0	2

The Shapiro-Wilk tests indicated that the data at baseline were normally distributed for most of the outcome measures. However, the test suggested the WOMAC stiffness score may not be normally distributed; $W(14) = 0.88$, $p=0.5$. Examination of the stem and leaf plot (figure-VIII.9) showed that this was primarily due to a particularly low score of one participant (TEN-13). The data was otherwise approximately normally distributed. Therefore, the data for all baseline outcome measures was considered to meet the assumptions required for the t-test (Field, 2009).

At baseline the means of the WOMAC pain score were 4.93 (SD 1.36) and 5.33 (SD 0.70) for the warm needle and control groups, respectively. There were no statistically significant differences between groups on any of the WOMAC scales or RAND-36 (v1.0) scores ([table-VII.2](#)). However, there were relatively large differences between the group means for the WOMAC function and the MHC. The MHC mean difference was 11.3 (SE 5.3) $t(12) = 2.1$, $p=0.06$. Suggesting that the mental health of the warm needle group was worse than the controls. Similarly, the WOMAC scale indicated that the warm needle group were less able to function: mean difference 1.30 (SE 0.75) $t(12) = 1.71$, $p=0.11$.

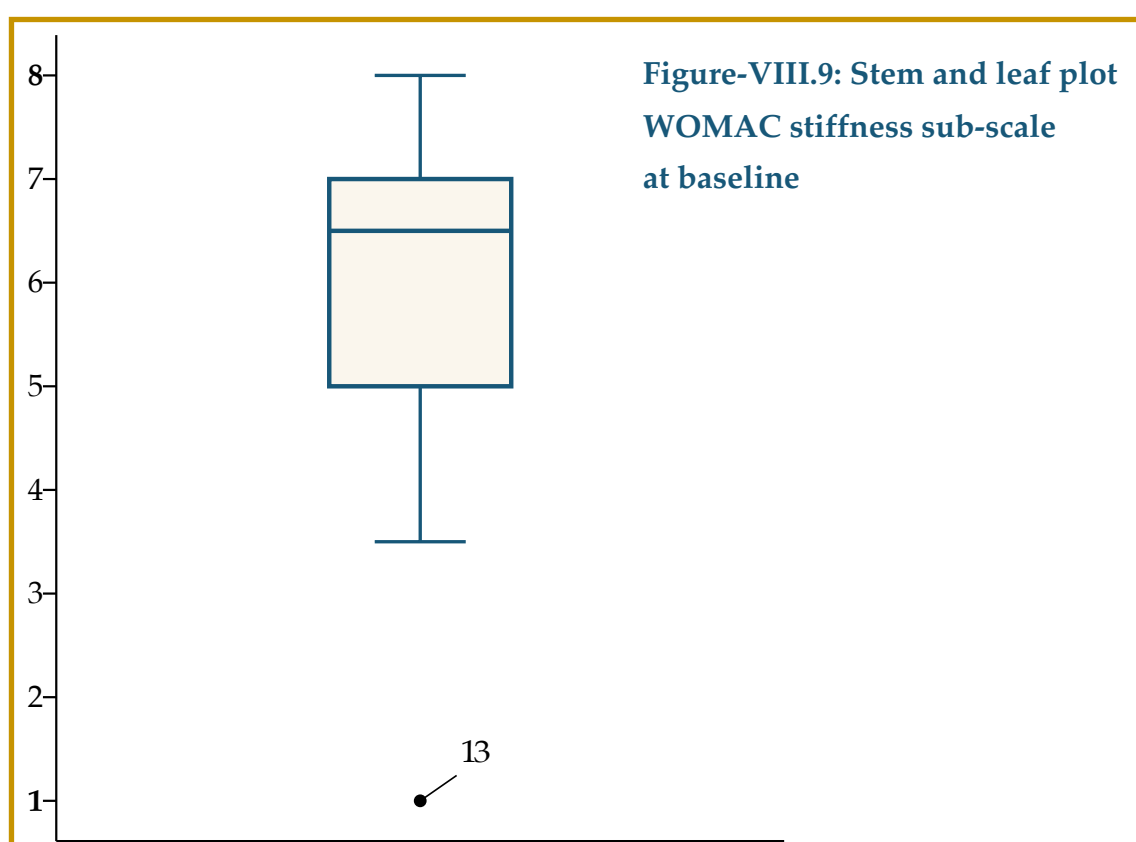


Table VII-2: WOMAC and RAND-36 (v1.0) scores at baseline

Time point (week)	Warm (n=8)		Needle (n=6)		t-test Equality of means			
	Mean	SD	Mean	SD	Mean Difference	SED	t	p value
WOMAC Pain	4.93	1.36	5.33	0.70	-0.41	0.61	-0.67	0.52
WOMAC Stiffness	5.81	1.71	6.00	2.51	-0.19	1.12	-0.17	0.87
WOMAC Function	4.98	1.46	3.69	1.31	1.30	0.75	1.71	0.11
WOMAC Total	5.24	1.06	5.01	1.04	0.20	0.57	0.41	0.69
RAND-36 (v1.0) PHC*	37.1	8.5	38.2	4.0	-5.7	8.4	-0.30	0.51
RAND-36 (v1.0) MHC*	35.3	11.0	46.6	8.0	-19.9	10.0	-2.12	0.07

SD - Standard deviation, SED Standard error of difference | Levene's test for equality of variances was non-significant for all tests - all tests 12 degrees of freedom
* a higher score indicates better health

VIII.3.3.Participant Interviews

Eight interviews were conducted: the shortest was 9 minutes, the longest 20 minutes; the mean 14.5 minutes. One participant provided written answers to the interview schedule. All interviews were conducted at LSBU. The interview schedule covered questions on what the participants felt about: the enrolment process; the information they were given; what motivated them; the disincentives; and whether the trial was as they had expected ([appendix-F.9](#)). The responses are summarised in [table-VIII.3](#).

All of those interviewed felt that the participant information booklet was comprehensive. The information, along with the opportunity to ask questions, ensured that participants were fully aware of the nature of the study. One interviewee thought that additional information regarding needle sensation could have been provided. Although, they acknowledged that they were particularly nervous about needles ([table-VIII.3](#)).

Most interviewees stated that their motivation was a combination of curiosity and personal need. Two people mentioned that biomedicine had not been helpful, another stated that they were drawn to the study because it wasn't a drug trial. One person had researched evidence for acupuncture and knee pain. For the majority, there were no disincentives to joining the study. One person was worried about the needles, another was concerned about the number of sessions and the amount of travel involved ([table-VIII.3](#)).

Table-VII.3: Summary of responses to questions regarding the recruitment process

TEN	Process	Information	Motivation	Disincentives	Met expectation
01	Straightforward	Very good	Curiosity and need	Time	No expectations
02	Fine	Very helpful	Need and interest	None	Yes
03	-	Clear	Interest - done trials before	None	Yes
04	Clear and fair	Easy to understand	Interest and need	None	Yes
07	Easy	Straightforward	Opportunity for treatment	None	Yes
08	Very thorough	Very thorough	Biomedicine no help	Needles	Yes
11	Very clear	Very thorough	Believe better than biomedicine	None	Did some reading so no need
13	-	Very thorough	Not a drug trial - researched evidence for acupuncture	Number of sessions and travel	-
14	Straightforward	-	Interest and seeking help	None	Yes

Table-VIII.4: Salient quotes from participants regarding the outcome measures

TEN	Usability	Comments	Suggestions
01	Some difficulty	Found it difficult to quantify, felt discussing the problem made more sense. Worried that other things happening in life would sway the results.	-
02	Easy	Wasn't sure why the questions on the emotions were included. However, also stated that the questionnaires highlighted how depressing the pain had become. Being told BMI lead to desire to make changes.	-
03	No problem	Interesting, helped to monitor progress	-
04	Easy	"Interesting, make me think more closely and constructively about my knee problems and other health issues"	Would be prepared to spend more time on them if necessary
07	Not difficult	"The BMI may me more conscious of getting healthier. Make me think about my physical needs and mobility. It was helpful and interesting"	-
08	Easy	"Wasn't sure why they focused so much on mental-health. Made me think about how things are connected if I have had a bad week seem to get more aches and pains"	-
11	Easy	"They were interesting, drew your attention to things. Time taken about right"	-
13	Not difficult	"I get frustrated when I can't answer in the way that I want to. Some questions were a little bit confusing for example conflating sitting and lying. They weren't desperately interesting. Timing about right"	Didn't allow me to differentiate pain on palpation and pain on movement
14	Straightforward	"Timing was fine. Confirmed how I felt"	

The interviewees did not appear to find the data collection process overly burdensome ([table-VIII.4](#)). One even stated they would be prepared to spend more time on them, if necessary. Generally, the outcome questionnaires were considered straightforward and easy to complete. TEN-13 felt two questions within the WOMAC that combined sitting and lying were problematic: question 4 “How much pain do you have sitting or lying?”; and question 7 “How severe is your stiffness after sitting, lying or resting later in the day?” For TEN-13 lying was much more painful than sitting. Another commented that they felt discussing problems made more sense, they felt quantitative measures had no real meaning. Several interviewees found completing the questionnaire interesting, and noted they became more aware of how they had been affected by the knee. Similarly, the body mass index (BMI) made a couple of interviewees more conscious of the need to get healthier. Two people said they were surprised by the questions regarding the emotions and mental health. Yet both, subsequently made comments about how ‘depressing the pain had become’ or that ‘if I’ve had a bad week I seem to get more aches and pains’.

Interviewees were also asked about the sensations that they felt during and immediately after acupuncture. The key points are presented in [table-VIII.5](#). A variety of adjectives were used to describe the needle sensations - tingling, deep ache, sting, faint discomfort. The common feature was that most interviewees felt a variety of sensations. Three people described sensations of warmth, two in the warm needle and one (TEN-08) in the needle acupuncture group. Someone in the warm needle group described ‘heat going down the routes of the meridians’. With the exception of TEN-14 all interviewees described feelings after the acupuncture. Typically these were full body feelings and were positive in nature; relaxed, calm, more ‘pep’ or light. Some people felt tired on occasions. Some people also describe specific local feelings “I walk back and my knees are sore for about an hour. It’s not soreness it’s just that they feel stiff and odd” (TEN-02). “The tingling [around the points], some of them went on into

the afternoon a couple of hours, and then it wore off” (TEN-03). Only TEN-14 had “no sort of sensation”.

The interviewees had a positive perception of acupuncture in general, at the end of their course of treatment ([table-VIII.6](#)). A couple highlighted acupuncture may not be useful for everyone. TEN-13 felt that ‘it didn't work for me’, TEN-08 had a relatively good outcome, but noted that it might not work for everyone. Many of the interviewees had never had acupuncture before, or had done so only briefly many years ago. Therefore, for most, it was difficult to make a judgement as to whether the treatment received in the study was similar to usual practice. TEN-01 and TEN-04 were both aware that treatment was focused on a specific area and there may have been greater variation, if the acupuncture was not been given as part of a trial ([table-VIII.6](#)).

Table-VII-5: Participant interviews - acupuncture effects

TEN	Needle sensations	After effects
01	Sometimes a quick sting. Sometimes nothing at all. Sometimes very deep ache running in different directions. Sensation of warmth	Fairly relaxed, and a couple of times quite weary
02	Sometimes sore, not painful but strange sensations going down your leg or burning or just funny sensations	My knees are sore for about an hour. It's different from any other pain I have in my knees. And then they feel fine.
03	Some hurt, some tingled, some you couldn't feel at all. Lots of different sensations	Sometimes the tingling went on into the afternoon a couple of hours, and then it wore off. — Felt revitalised
04	Sometimes painful, other times painless. On a couple occasions I needed the needles to be taken out because of the discomfort.	Immediately afterwards, I felt a little tired. Then had more energy than before. The needles did not have an instant effect. But I felt the changes later
07	Usually no feeling. Occasionally painful when inserted into painful areas, but that reduced.	Quite calm and relaxed
08	Generally no pain. Slight tingling or warmth	Felt more 'pep' an hour to an hour and half later
11	Like electricity feeling, it's hard to explain. And heat going down the routes of the meridians	Like I was walking on clouds, my legs feel so light and new.
13	Sharp pain when inserted. A couple of times a dull ache around the needles. Others just sort of faint discomfort	A little light, I don't mean light headed, its quite a nice feeling. It doesn't last long but it's quite pleasant
14	Warmth and a bit of pain around them	No sort of sensation

Control group highlighted in green

Table-VII.6: Attitudes to acupuncture and perception of how the study differed from acupuncture in practice

TEN	Acupuncture in general	Study acupuncture compared to usual practice
01	Fascinating, I want it to work, I'm not 100% sure that it does; other than because people believe it does.	I never have had acupuncture on a regular basis constantly working on one area.
02	-	I wouldn't know really because I have never had the experience before
03	Now, I would sing its praises	Don't know because had never had acupuncture before
04	I think acupuncture has a proven place in medicine. So, my attitude is a positive one. But that doesn't mean that I expect it to always solve every medical problem.	There would probably be variations in my treatment otherwise, as I reacted positively or negatively to the needling.
07	It was good, I'm confident about it. But I was confident in the beginning anyway, because of my experience 30 some years ago	Thought similar to usual practice
08	I would recommend it to others, I think it is useful thing, I don't think it works for everyone but I certainly think it would be of benefit to a large number of people.	Difficult to say as never had acupuncture before
11	I do trust in acupuncture, and I think it's a great additional help, which one can count on - except the tablets. So definitely positive	As I have never had an experience, but I think this is how it should be. This is what I read how it looks like all feels like. I don't think that there was anything different
13	I think I probably wouldn't try it again. I think three out of four not getting a result, suggests that for me it doesn't work. But then neither the most drugs work for most people.	Normally there would be fewer needles and less pain. Otherwise not really very different, no.
14	-	I would hope that I would have similar good treatment, if I went to an outside practitioner.

VIII.3.4.Researcher's reflection on the pilot study

As the researcher and the acupuncturist, it was impossible to extricate myself from numerous decisions that I made during this study. In addition, I thought it was important to reflect on my experiences. Therefore, this section has been written in the first person.

I chose the condition of osteoarthritis as I believed that a semi-flexible protocol would be able to mirror usual practice to a reasonable degree. And I still believe that to be the case. Nevertheless, at times the protocol felt more restrictive than I had anticipated, even though I had designed it myself. This was most apparent when treating TEN-04 and TEN-14. TEN-04 was very sensitive to needling. This could be dealt with within the protocol because it was not mandated to stimulate the needles to achieve the deqi. Which I think for this particular person would've had a negative effect. Techniques such as, slightly reducing the number of needles, taking care to ensure the participant breathed out at the time of needle insertion, enabled treatment to continue. I began to feel that the problem was fundamentally one of stress and not really knee pain. In usual practice I could've tried more distal points away from knee. The objective, would have been to relax the patient and allow them to become more comfortable with the process of acupuncture. This change of approach may or may not have happened if I had seen TEN-04 in a normal clinic setting. TEN-04 appeared to be responding reasonably well at the beginning of the treatment which encouraged me to continue with the protocol. However, I was aware that I had a conflict of interest. On the one hand, I wished the person to continue within the trial, plus I did not know a change of approach would be more effective. On the other hand, the change of approach might have helped. In usual practice I would be free to make my decision, without the additional consideration of successfully collecting data.

TEN-14 did not respond very well to the acupuncture treatment, they were the only participant whose condition got worse. It became clear relatively early on that there were numerous factors, beyond the knee pain, impacting on their

well-being. In particular they were confronted with a particularly stressful situation related to their living environment. TEN-14 on several occasions was clearly upset by the ongoing difficulties that they faced. It was combining with the knee pain to prevent them from sleeping. As with TEN-04 there was no guarantee that a change of approach would have made the acupuncture more effective, and I am not particularly confident it would have done. Nevertheless, I would probably have made some changes in usual practice. This is part of acupuncture practice, if the patient does not respond then one naturally tries a different approach.

Both of these situations highlighted an inherent ethical problem of using restricted acupuncture protocols in clinical trials. For many acupuncturists it is inevitable that they will feel on occasions the acupuncture should be done differently. This will occur even if a condition is chosen specifically because it is thought that a semi-flexible protocol would largely mirror usual practice; as was the case in this study. However, no practitioner can be sure that a different approach would be more beneficial unless research has already been conducted.

TEN-02 surprised me at the mid-point data collection. I had felt that the treatment was going well, from the feedback I was getting at each treatment. Yet, when I looked at the WOMAC pain score I realised there had been virtually no change. It later transpired that TEN-02 had actually played tennis a couple of days prior to completing the form. At the end of treatment data collection point the pain score dropped to zero. What surprised me most, was the level of activity. Other participants with similar pain scores appeared to be affected by their knee pain to such a degree that tasks far less strenuous than playing tennis were challenging. For example, TEN-05 who had same baseline pain score had obvious difficulties in walking.

I feel that this experience of TEN-02 has given me a much greater appreciation of the difficulties in assessing outcomes. The pain scale was part of the eligibility criteria. Because of this, I feel it had a greater level of importance in

my mind. However, looked at in isolation it is clear that improvement, the ability of TEN-02 to be more active, was not captured by a single pain scale. In addition, I have a greater understanding of the subjectivity of patient centred outcome measures. I had of course known that the outcome measures were subjective prior to the study. Seeing two people with the same pain score who appear to have very different levels of disability, from my perspective, underlined the subjectivity of the outcomes.

I was surprised how difficult it was not to be hyper-vigilant regarding group allocation of participants. I found myself looking for various clues, such as ash on the floor, and considering these indicated the person was in the warm needle. An assistant mentioned that he thought one of the participants was receiving warm needle at some treatments but not at others. Therefore, after the first block of 10, the used envelopes were examined. There was no problem with the allocation process, the assistant had been mistaken. At that point, I tested myself. I had guessed correctly only 4/10 times. I do believe the process can successfully blind acupuncturists. On reflection, however, the study should have been designed with a more formal process of assessing acupuncturist blinding. For example, the acupuncturist could guess which group each participant had been assigned to at a certain time point.

Generally, I believe the recruitment process went well and the eligibility criteria were adequate. There were two participants who I felt were borderline cases. TEN-10 had been diagnosed as having osteoarthritis due to an earlier injury. TEN-10 was younger than the other participants (30 years old), as a consequence, they did not really fit the profile I was expecting. There was no lower age limit, so he did meet the criteria. In terms of TCM his knee problem was suitable. I had already accepted another participant who had also been told the osteoarthritis was due to an earlier injury. They were in their late 40s and fitted my expectations more completely. I decided I should stick strictly to the inclusion criteria as had been set out in protocol, so TEN-10 was enrolled.

TEN-13 took analgesics, when required, for headaches. At baseline I understood this to be below the threshold that would have excluded them from the trial. At the end of treatment data collection point, TEN-13 reported that they had needed to take the headache medication twice within the week. Therefore they may have exceeded the limit of three times in four weeks (table-VIII.9). The usage appeared to be on the borderline rather than a clear breach, therefore TEN-13's data was included in the analysis.

VIII.3.5.Evaluation of blinding procedure

Participants were asked a direct question regarding group allocation at the mid-point and the end of treatment, then asked to give a reason for their choice (section-VIII.2.1.d). The responses are presented in table-VIII.7. The warm needle group are in the top half the table and the control group in the bottom half (highlighted in green).

Most participants in the warm group correctly identified their allocation. Participants appear to have been able to sense the heat on some occasions. The most extreme example was TEN-01 who felt the heat unambiguously during one of the early sessions, to the extent that they considered asking for the moxa to be removed. At the time it was surmised that the angle of needle insertion had left a patch of skin exposed to direct radiated heat. From that point, larger skin guards were used and additional guards could be added, if it was felt some skin was exposed. Such a strong degree of heat did not appear to be felt by other participants. TEN-05 was not given warm needle acupuncture on one occasion because they were feeling particularly nervous, they were informed that the moxa would not be lit. They cited the different sensation on this occasion as the reason for believing they were usually receiving warm needle acupuncture.

It is not clear why some of the participants could feel the heat sometimes. It may have been due to: small gaps between patches, convection leading to the heat rising above the guards, radiated heat through the hole in the skin guard,

sensing the heat conducted along the needle. The participants did not feel the heat at each treatment and TEN-04 did not guess that they were in the warm needle group. The weakness of this method was that the guards only need to be inadequately arranged on one occasion for blinding to be broken.

Only one person guessed they were in the needle acupuncture group, and they were correct. The other five participants in the needle acupuncture group thought they were having warm needle acupuncture, although four also said 'don't know' at the other data collection point ([table-VIII.7](#)). The needle acupuncture group were often less certain in their responses. TEN-06 selected 'warm', but then added the caveat 'I don't know if it is in my head', TEN-08 said they were 60% sure ([table-VIII.7](#)).

Participants in both groups appear to have been influenced by the smell of the moxa and the actions of the assistants. TEN-07 thought that the assistants were not 'lighting the moxa as much as in beginning'. TEN-14 and TEN-10 both indicated that the smell led them to believe they were in the moxa the group, although they were in different groups. The sensation of the needles being touched suggested to TEN-13 that they were having warm needle acupuncture, but they were not.

Table-VII.7: Participant guesses to group allocation questions and free text explanations

TEN	Mid-point	End of Treatment	Mid-point	End of Treatment
1	Warm	Warm	One needle felt hot in one of the first session. "Practitioner acting, not so good"	I had formed an idea
2	Warm	Warm	"The way that the moxa lighters do the lighting"	Could feel absolute heat the week before last. My knee jerked - this was just a twitch
4	Don't know	Don't know	na	na
5	Warm	Warm	Sensation of heat - noticed the difference from the time when knew no heat was applied	"It is how it feels compared to the time I didn't have moxa"
9	Warm	Warm	Some treatments the needles felt warm. Only some of the needles	Couple of sessions could feel the heat
12	Warm	Warm	Could feel the heat sometimes	Feeling it hot sometimes
14	Don't know	Warm	na	The smell, did feel at time my knees were warmer
3	Warm	Don't know	Instinct	na
6	Warm	Don't know	I think its warm - don't know if it is in my head	More inclined to say without now. It felt warm in the beginning
7	Needle	Needle	Moxa lighters don't seem to be lighting the moxa as much as in the beginning	Didn't smell anything
8	Don't know	Warm	na	Not 100% sure. It is the smell in the room, 60% sure
10	Don't know	Warm	na	I can smell something
13	Warm	Warm	I can feel the needles being touched	Felt needles being touched

Control group highlighted in green

Statistical evaluation of the success of blinding

The responses at midpoint and end of treatment were combined. This assumed that the guesses of individual participants at each time point were independent events. If blinding was successful, this should be the case. The responses of TEN-01 and TEN-05 were removed because they were considered to have been unblinded by avoidable failures of process (see above). The indexes were:

$$BI_{\text{warm}} = 0.75 \text{ (CI } 0.5 - 1) \quad BI_{\text{needle}} = -0.33 \text{ (CI } -0.75 - 0.12)$$

Bang et al suggested that unblinding can be claimed if the confidence interval does not cross 0, which was the case in the warm needle group (Bang et al., 2010). Blinding appears to have been successful for the needle acupuncture group.

Overall evaluation of the blinding procedures

The statistical assessment and the qualitative information indicated that blinding was not successful in the warm needle group, however, the control group also generally thought they were having warm needle acupuncture. It has been suggested that participants tend to have an optimistic attitude and assume they are in the experimental group (Bang et al., 2010). In the control group 5/6 opted at some point for 'warm', only one participant ever guessed 'needle' and they cited changes in the behaviour of the assistants as an explanation. It seems reasonable to assume that random guessing Bang $BI = 0$ was unlikely to be achieved. The objective of blinding is to control for the psychological effects on the outcomes. This could be achieved if both groups generally believe they are having warm needle acupuncture.

VIII.3.6.Expectation

Fisher's exact test was used to assess whether there was any relationship between expectation and outcomes (section-VII.2.2). It was assumed that, if expectations were to influence outcomes, then this would be most apparent at the first data collection point after baseline, mid-point. No relationship between responder rates and expectations was found (table-VIII.8). The only question

that showed a trend towards a relationship was “How confident that you feel that this treatment can alleviate your knee pain?” . This was asked at the fourth treatment session. At the fourth treatment more participants had high expectations of what acupuncture could do for them personally, 10/14 compared to 5/14 at baseline. This would suggest that the relationship was due to early treatment success rather than preconceived expectations of acupuncture. A secondary analysis was carried out using the responder rates at the end of treatment - no relationship was found (appendix-F.11).

Table-VII.8: Expectations and mid-point pain - Fisher’s exact test

How effective do you consider acupuncture in general?			
	Low	High	Fisher Sig 1.00
Non-responder	3	5	
Responder	2	4	
What do you personally expect from the acupuncture treatment you will receive?			
	Low	High	Fisher Sig 1.00
Non-responder	5	3	
Responder	4	2	
Do you think warm needle acupuncture will be more effective than needle acupuncture for your knees?			
	Low	High	Fisher Sig 1.00
Non-responder	4	4	
Responder	3	3	
How confident do you feel that this treatment can alleviate your knee pain?			
	Low	High	Fisher Sig 0.09
Non-responder	4	4	
Responder	0	6	

2 sided significance

VIII.3.7.Medication

Participants were asked about their medication usage ([section-VII.2.1.e](#)), at each of the four time-points. At baseline six participants were not using any pain medication. The most commonly used analgesics were paracetamol and ibuprofen ([table-VIII.9](#)). Of the eight participants who had used medication in the week before baseline, three did not need to do so every day. These were considered to have low usage. Five had taken more than seven tablets in the week, and were considered to have high usage. Four of these were able to substantially reduce their use of medication by end of treatment.

A number of participants were sent the follow-up data collection forms in the post for completion. Some of these left the medication section blank. Therefore, it was unclear whether this section had simply been missed or they had not taken any medication. Ambiguity also arose when not all medications recorded at baseline were identified at subsequent time-points. For example, at baseline TEN-13 was recorded as having taken sumatriptan and paracetamol, at the midpoint only sumatriptan was recorded. This meant it was unclear whether paracetamol had not been used or had been overlooked.

Table-VIII.9: Medication usage

Participant	Medication	Week 0	Week 4	Week 8	Week 16
TEN 1 †	Paracetamol	2/w	None	None	Missing
	Ibuprofen	-	1/d	None	
TEN 2		None	None	None	None
TEN 3 †	Paracetamol	2/w	Missing	None	None
TEN 4		None	None	Missing	Missing
TEN 5*	Ibuprofen 400mg	12/w	None	2/w	Missing
	Solpadol 30mg	12/w	None	None	Missing
TEN 6		None	None	None	Missing
TEN 7		None	None	None	Missing
TEN 8*	Ibuprofen 200mg	28/w	4/w	None	Unclear
TEN 9*	Ibuprofen 400mg	2/w	250mg 2/w	None	Missing
	Paracetamol 200mg	9/w	3/w	None	Missing
TEN 10		None	None	None	None
TEN 11*	Ibuprofen	30/w	3/w	2/w	None
	Paracetamol	30/w	3/w	None	None
TEN 12		None	Missing	None	None
TEN 13 †	Paracetamol 200mg	2/w	Unclear	1/w	Unclear
	Ibuprofen 200mg	-	-	1/w	Unclear
	Sumatriptan 50mg	1/w	1/w	2/w	Unclear
TEN 14**	Ibuprofen 200mg	14/w	16/w	28/w	28/w
	Lansoprazole 30mg	7/w	Unclear	Unclear	7/w
	Amitriptyline 10mg				7/w

† low usage, * high usage with reduction, ** high usage no reduction

VIII.3.8.Safety

None of the participants reported severe or long lasting adverse effects. The majority reported bruising at some stage during their course of treatment. There were no adverse events specifically related to the use of moxibustion. Comments from each of the participants who were interviewed are presented in [table-VIII.10](#).

Apart from bruising, there were two incidents. TEN-03 felt faint during their first treatment session. The needles were removed and the participant lay down on the couch and shortly recovered. At the time TEN-03 stated that they thought the room had been too warm; and this was the cause of feeling faint rather than the acupuncture. No further treatment was given on that day, however, TEN-03 was happy to continue in the study. There was no recurrence of feeling faint during the subsequent 11 treatments. The acupuncturists and assistants paid particular attention to ensuring the treatment room was a comfortable temperature.

TEN-05 arrived at the clinic for treatment five feeling unsettled. Once the needles were inserted they started to sweat from their palms. TEN-05 was happy to sit with the needles inserted for the normal treatment time, but it was agreed that moxa would not be lit. This would ensure the needles could be quickly removed, if TEN-05 began to feel uncomfortable.

Table-VII.10: Interviewees comment on adverse events

TEN	Sentence
01	No negative effects
02	None
03	Just the dizziness, apart from that it was all positive
04	I have not noticed any negative effects from having acupuncture. Sometimes it has been painful, but that feeling goes away quickly.
07	No harms
08	Only slight cosmetic ones, as I say, I had bruising once or twice.
11	None
13	No harms
14	No negative effects

VIII.3.9. Pattern differentiation

The bi-syndrome differential patterns diagnosed at baseline are presented in [table-VIII.1](#). Due to the small number of participants this data was not used in any analysis.

VIII.3.10. Outcomes measures WOMAC and RAND-36

Completed questionnaires for both WOMAC and RAND-36 (v1.0) were obtained for all participants at midpoint and end of treatment. Two participants did not complete the questionnaires at follow-up, both were in the warm needle group ([section-VII.3.1](#)). The data from the last observation was carried forward (week 8), which is common practice for intention-to-treat (ITT) analysis (White et al., 2011).

VIII.3.10.a. Change over time - repeated measures

The WOMAC scores showed an improvement in all three dimensions, and the global score, from baseline to weeks 4, 8 and 16 ([appendix-F.12](#)). When considering all 14 participants, at end of treatment and follow-up the changes were statistically significant ([appendix-F.12](#)). For example, the baseline to follow-up mean change for pain was -2.7 (SE 0.50), $t(13) = 5.43$, $p < 0.001$, $r=0.8$. The same pattern of improvement was observed when the groups were considered separately, although statistical significance was not always reached ([table-VIII.11](#)) ([appendix-F.12](#)). This was probably due to the decrease in statistical power.

The Rand-36 (v1.0) physical health composite (PHC) showed an improvement, that was statistically significant at end of treatment, when considering all 14 participants mean change was 5.4 (SE 2.2), $t(13) = 2.47$, $p=0.03$, $r=0.6$ ([appendix-F.13](#)). At follow-up the PHC change was no longer statistically significant, 4.7 (SE 2.23), $t(13) = 2.10$, $p=0.06$, $r=0.5$. The mental health composite (MHC) also improved across the same time period, but this was not statistically significant at either time-point ([appendix-F.13](#)). At follow-up the mean change was 4.91 (SE 3.78), $t(13) = 1.3$, $p=0.22$, $r=0.34$.

Table-VIII.11: WOMAC and RAND-36 at the four time-points

Time point (week)	Warm		Needle	
Baseline (0)	Mean	SEM	Mean	SEM
WOMAC Pain	4.93	0.48	5.33	0.29
WOMAC Stiffness	5.81	0.60	6.00	1.02
WOMAC Function	4.98	0.52	3.69	0.53
WOMAC Total	5.24	0.38	5.01	0.43
RAND-36 PHC	37.1	3.02	38.23	1.65
MHC-36 MHC	35.3	3.88	46.59	3.27
Mid-point (4)	Mean	SEM	Mean	SEM
WOMAC Pain	3.65 [†]	0.80	3.27 [†]	0.61
WOMAC Stiffness	4.06	0.86	5.00	0.97
WOMAC Function	3.24 [†]	0.48	2.97	0.76
WOMAC Total	3.65 [†]	0.67	3.75	0.70
RAND-36 PHC	41.92	3.36	40.76	1.93
MHC-36 MHC	41.45	3.92	47.90	2.91
End of treatment (8)	Mean	SEM	Mean	SEM
WOMAC Pain	3.38	1.06	1.78 [†]	0.72
WOMAC Stiffness	3.94	1.00	2.92 [†]	0.71
WOMAC Function	2.59 [†]	0.90	1.75	0.70
WOMAC Total	3.30	0.97	2.15 [†]	0.63
RAND-36 PHC	41.72	4.23	44.69	3.53
MHC-36 MHC	41.14	5.76	47.27	4.77
Follow-up (16)	Mean	SEM	Mean	SEM
WOMAC Pain	2.25 [†]	0.66	2.60 [†]	0.75
WOMAC Stiffness	2.69 [†]	0.63	3.42	0.75
WOMAC Function	2.28 [†]	0.90	1.65	0.63
WOMAC Total	2.40 [†]	0.70	2.55 [†]	0.59
RAND-36 PHC	40.28	4.14	44.97	3.74
MHC-36 MHC	42.30	5.29	48.69	4.51

[†] significant change from baseline $p < 0.5$

Paired t-tests, all 14 participants, comparing the change in WOMAC from mid point to follow-up showed statistically significant differences for all three sub-scales and the global score, $p < 0.05$ (appendix-F.12.d). This suggests that the

five subsequent treatments were beneficial, and that a shorter course of acupuncture may be inadequate.

VIII.3.10.b. Inter-group comparisons

Independent t-tests found no statistically significant differences between the groups for the global WOMAC score ([table-VIII.12](#)) or any of the dimensions, at the end of treatment nor at the follow-up ([appendix-F.14](#)). No significant differences were seen for the PHC and MHC scores ([appendix-F.15](#)). As can be seen from [figure-VIII.10/11](#) the confidence intervals were very wide, illustrating that the study was too small to detect inter-group differences.

Table-VIII.12: Independent t-tests WOMAC global score

Time point	Levene's Test		t-test for Equality of Means			
	F	Sig.	t	p	MD	SED
Mid-point	1.15	0.30	0.73	0.48	0.63	0.86
End of treatment	0.11	0.74	-0.14	0.89	-0.20	1.40
Follow-up	0.44	0.52	0.39	0.70	0.50	1.29

df = 12, MD - mean difference (Warm needle - control), SED - Standard Error Difference

Figure-VIII.10: Changes in the WOMAC Index over time

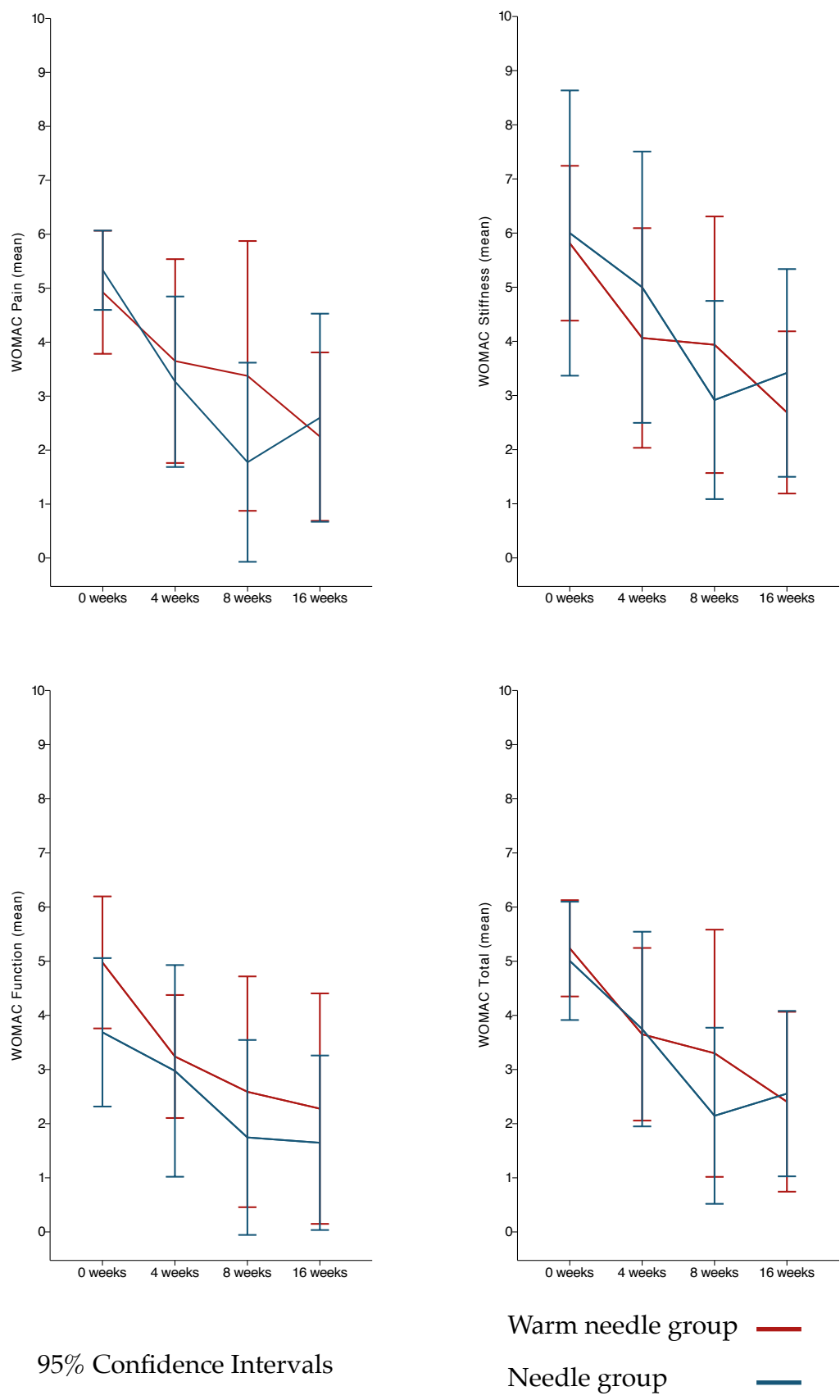
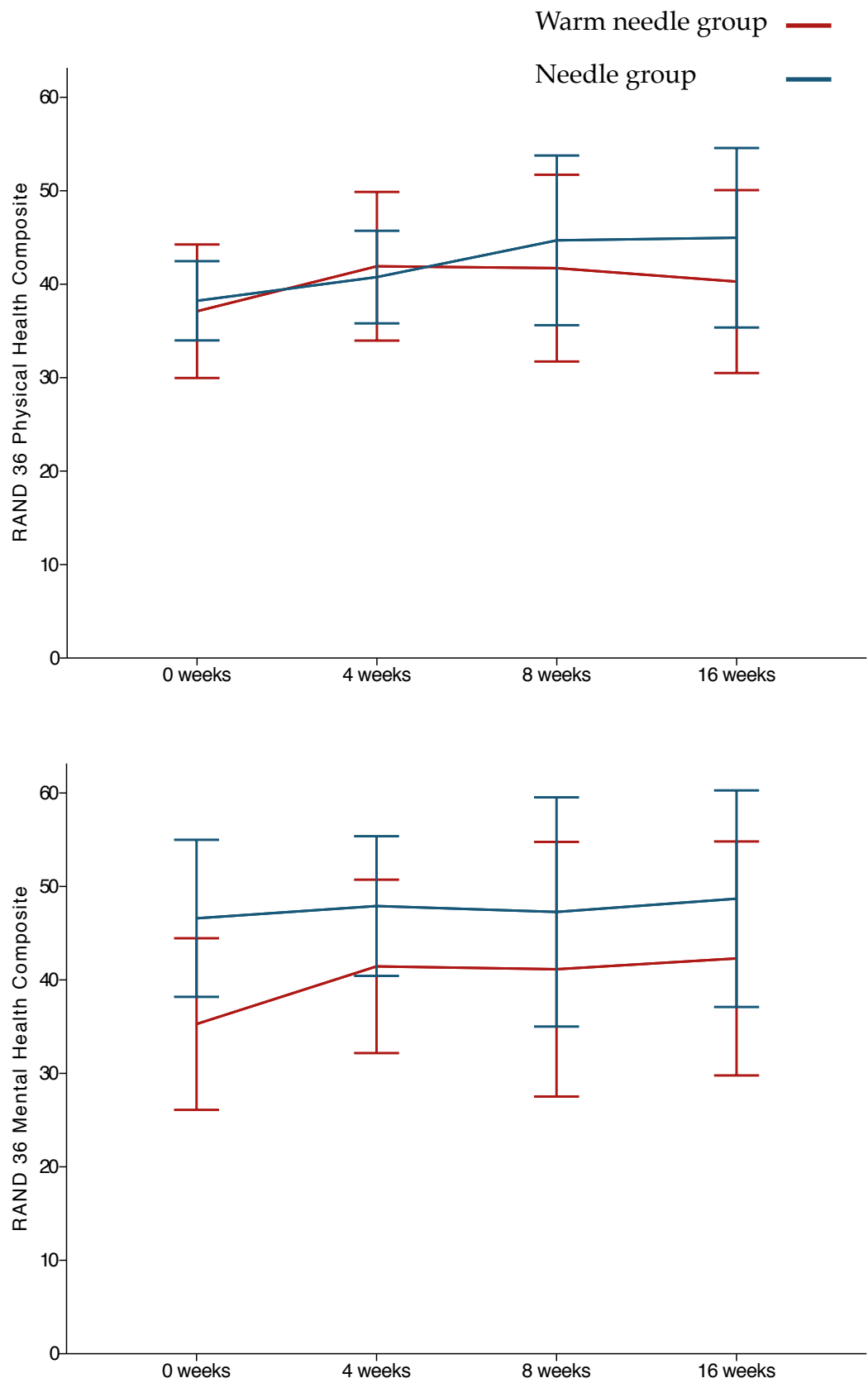


Figure-VIII.11: Changes in RAND-36 PHC and MHC over time



VIII.3.10.c.Responder rates

Responder rates have been recommended as a means of identifying subsets of patients who receive clinically meaningful benefits (Alper et al., 2016). A 50% improvement from baseline has previously been used as a criterion for identifying responders (Vickers et al., 2012). At the end of treatment 3/8 and 4/6 were responders in the warm needle and control groups, respectively. At follow-up 5/8 in the warm needle group were responders, and 3/6 in the needle acupuncture group. The percentage change in the WOMAC global score from baseline has been illustrated in [figure-VIII.12](#), with a horizontal dotted line marking the 50% criterion.

There was a difference at baseline between the two groups for the RAND-36 measure of mental health (MHC) that approached statistical significance ([section-VII.3.2](#)). The individual scores have been illustrated in [figure-VIII.13](#). The Rand-36 (v1.0) composites are linear T-scores, which have a mean of 50 and standard deviation of 10 (Hays et al., 1993). We can expect the MHC to have a mean of approximately 43 in those who are depressed (Hays et al., 1993). There were three individuals who consistently scored below 43, all of them were in the warm needle group, TEN 4, 9 and 14. These were the three non-responders ([figure-VII.13](#)).

Figure-VIII.12: Percentage changes in WOMAC global scores for individual participants

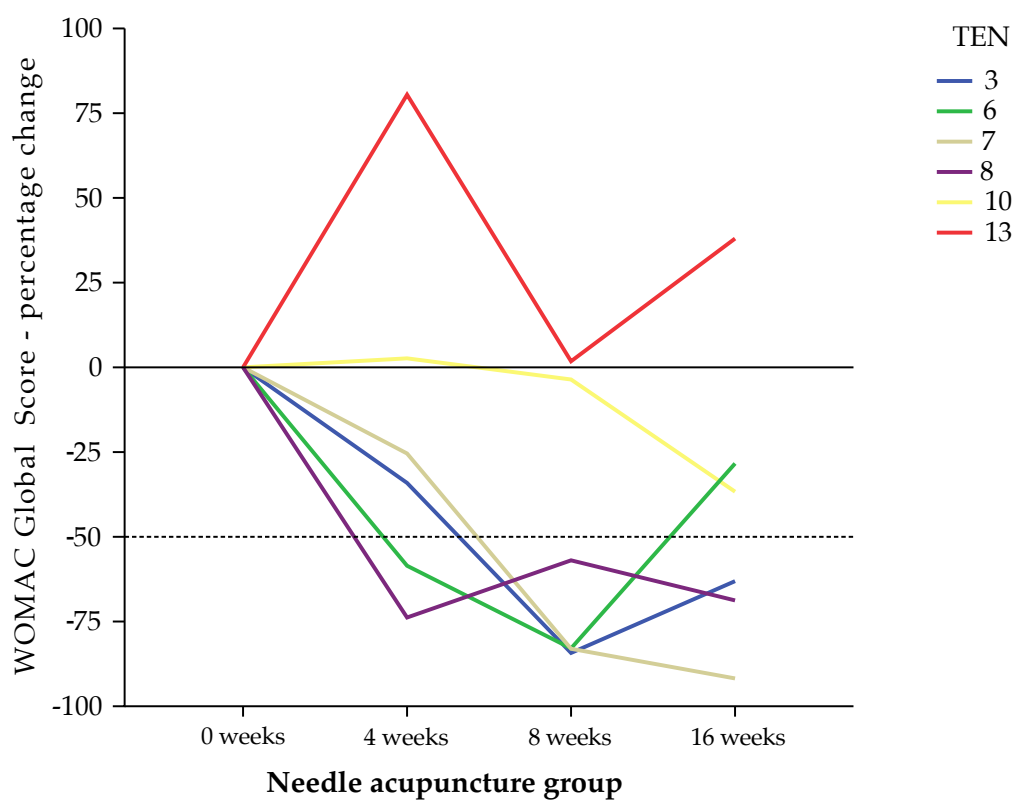
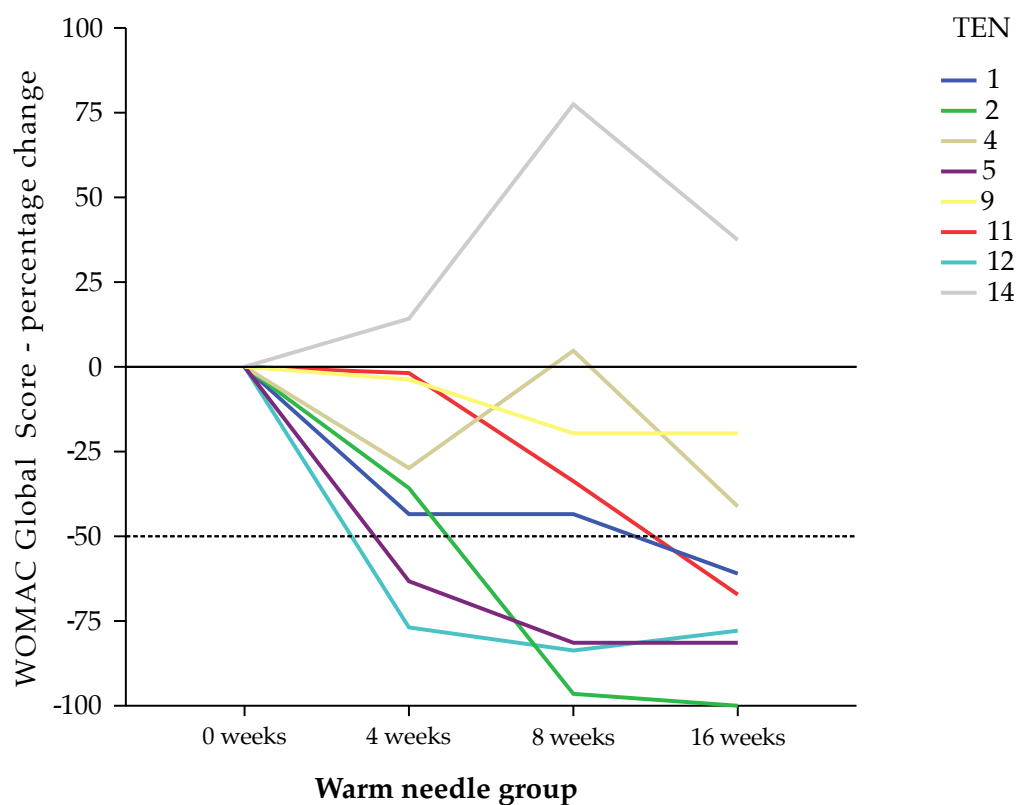
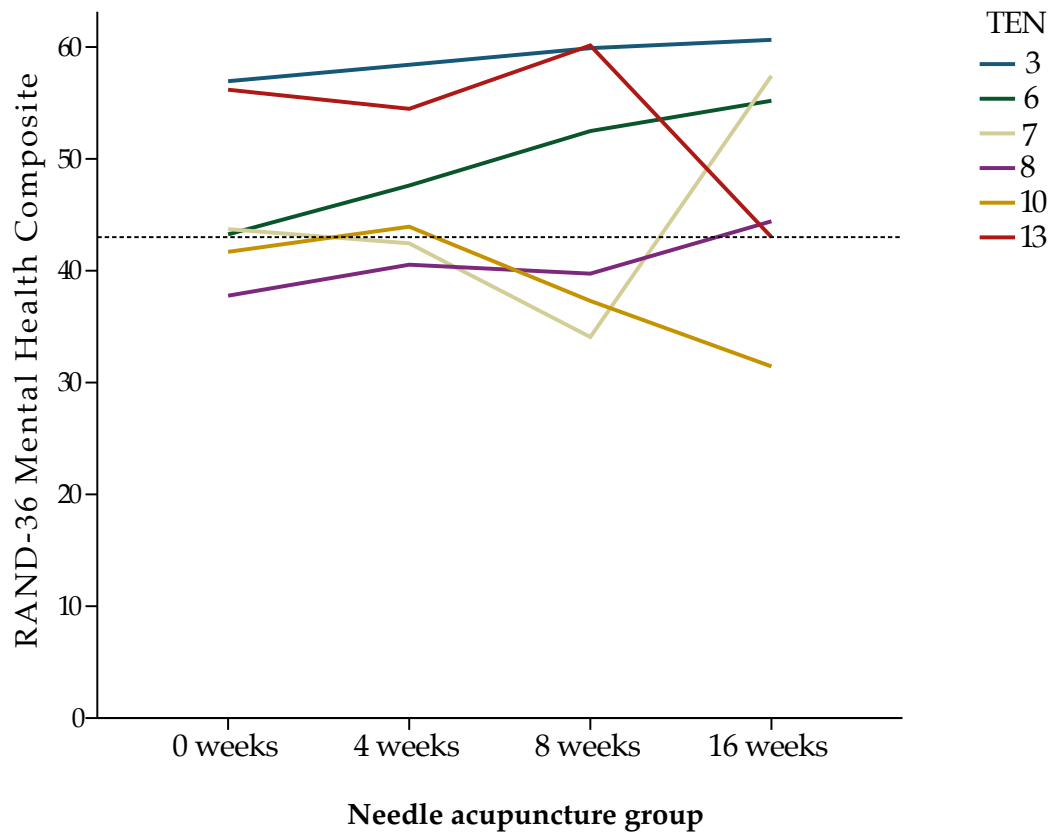
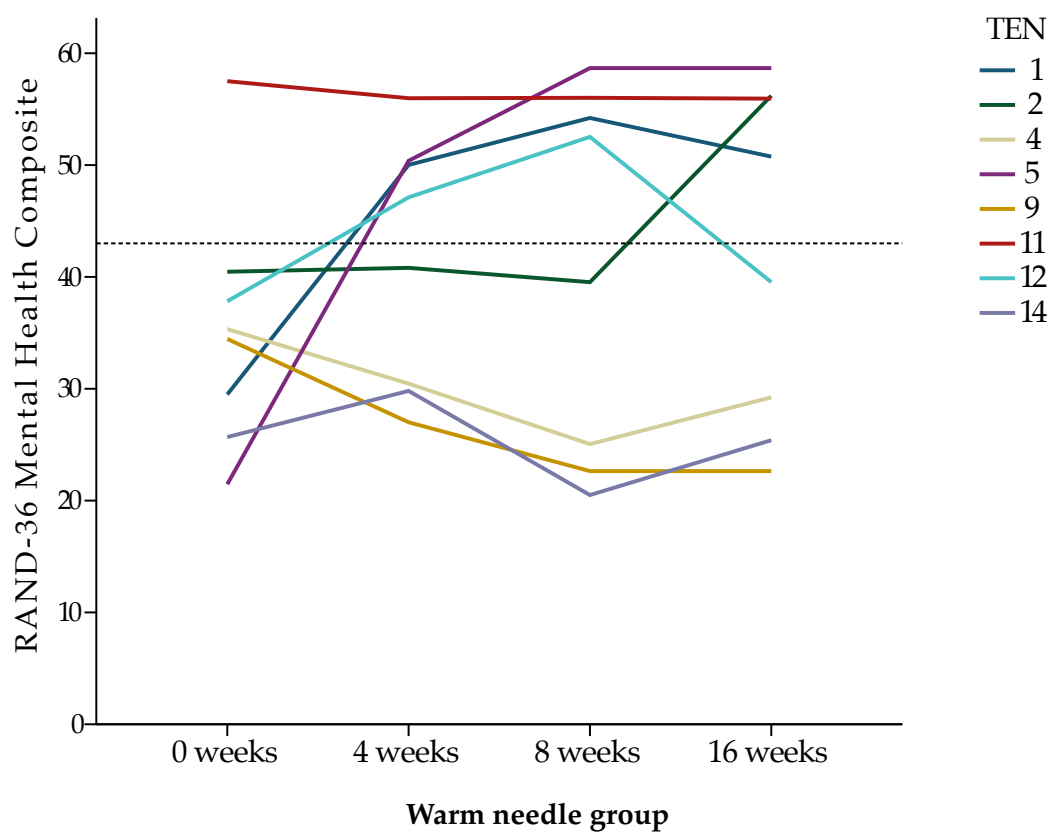


Figure-VIII-13: RAND-36 Individual MHC scores line graph



VIII.3.10.d. Issues of concern with the outcome measures

The issues that arose in the administration and interpretation of the WOMAC Index centred around the stiffness dimension. There were only two questions that related to stiffness, compare to five for pain, and 17 for function. There was a comparatively higher degree of variance compared to the other two dimensions. For example, the standard deviation for the mean change between weeks 0 and 8 was 3.27 for stiffness, compared to 2.56 and 2.58 for function and pain, respectively. TEN-13 stated that they felt it was difficult to answer questions that conflated sitting and lying ([section-VII.3.3](#)) this included one of the stiffness questions (Q7): How severe is your stiffness after sitting, lying or resting later in the day? TEN-13's score for stiffness was 1.00 at baseline then 6.50 at the midpoint; 2.50 and 4.00 at weeks 8 and 16, respectively. It seems likely that some of this variation was due to different interpretations of questions rather than actual changes in the disease condition.

The RAND-36 measures eight components of health ([section-VII.2.2](#)), these were initially calculated. However some of these sub-scales had too many ceiling and basement scores. For example, at baseline for role limitation due to emotional problems, eight participants scored 100 and three participants scored 0. It was felt that analysis of these sub-scales would not be able to detect change and it would be better to use the two composite scores, PHC and MHC.

VIII.4. Discussion

This study has shown the participant information booklet, consent form and recruitment procedures were effective in preparing the potential participants prior to entering the study. The participants appear to have been fully aware of the commitments, the burden of data collection, and the nature of the treatments. The randomisation and allocation procedures also worked smoothly.

There were no major issues with the eligibility criteria. It proved difficult to obtain documentary evidence of an osteoarthritis diagnosis for each participant. The researcher is not a medical doctor, and so is not appropriately qualified to make a formal diagnosis. Unfortunately, it had not been possible to gain support from the local health trust in sourcing potential participants. Resources were also not available for radiographic scans. Participants were asked to provide evidence of diagnoses, however, not all participants had kept this information. The lack of radiographic evidence or a formal diagnosis of osteoarthritis, in some cases, was not thought to affect the conclusions of this study. The focus of this study was on testing the procedures not the hypothesis. The protocol was based on TCM and the researcher is qualified to make a TCM diagnosis. A full-scale RCT should aim to produce results that are meaningful to the wider scientific community. Therefore, a diagnosis by a properly qualified individual and radiographic evidence should be obtained.

The blinding procedures were not 100% effective, as the warm needle group were generally more certain of their group allocation. The needle acupuncture group were to some extent deceived into believing they were also having warm needle acupuncture. Therefore, the procedures were partially successful. It should be possible to improve the blinding procedures. As participants in clinical trials have a tendency to choose the 'active' treatment, it is unlikely that group allocation guesses would ever show a 50-50 ratio (Cheon et al., 2018).

In essence, the objective is to make those in the warm needle group uncertain, and convince the control group that they are having moxibustion. Greater care in positioning the skin guards may reduce the instances of participants sensing the radiated heat. Training should emphasise to assistants that they must enact the use of moxibustion at each session. In a full-scale trial a supervisor could be employed to monitor the assistants and to check the positioning of the skin guards. A reduction in the number of treatments would also reduce the potential for allocation to be revealed due to poorly positioned skin guards.

However, as the five post mid-point treatments led to a further reduction in symptoms this may not be appropriate.

The information given to participants prior to enrolment could be altered to imply that they would receive warm needle on some occasions and needle acupuncture on others. In reality they would remain randomised to two distinct groups. If the patches were incorrectly applied on one occasion, then blinding would only be broken for that treatment. The participant would not know what kind of acupuncture they would receive at the next session. At each session participants would be asked which treatment they thought they had received. Consideration would need to be given regarding the ethics of deceiving participants in this way.

Warm needle acupuncture appears to be acceptable to UK patients. None of the participants were perturbed by being unable to see the needles and moxa during the treatment. There were no adverse events beyond those that are known to be associated with acupuncture (MacPherson et al., 2001, Witt et al., 2009). However, sitting is a less frequently used position for acupuncture. It is important that the chair is positioned close to a couch, incase the patient begins to feel faint.

It was not possible to use the results of this study for a power calculation. The cohorts were too small for any meaningful comparison between groups. In addition the potential imbalances between the groups, lower MHC and higher WOMAC function score in the warm needle group, may have affected the results. The study was also too small to utilise the TCM pattern differentiation data to assess whether those identified with a particular pattern were more or less likely to respond. Nevertheless in retrospect the collection methods for the pattern differentiation data could be improved. A system whereby more than one practitioner assesses each participant should be included. This would enable inter-rater comparisons to be made and ensure greater reliability of the study.

The data has been analysed using both comparison of means and as responder rates. Pain has been identified as the primary dimension rather than stiffness or function, in previous research (Vickers et al., 2012). Of the three WOMAC subscales pain is thought to be most sensitive to change and function the least sensitive (Bellamy, 2015). However, relying too much on the pain scale may lead to a narrow interpretation. There may be a complex journey of improvement that is different between patients. For example, TEN-02 at the mid-point had only had a 21% reduction in pain. However, the function had improved by 51%. At follow-up TEN-02 was pain free. Improvement in function may have knock-on therapeutic effects as the individual is able to exercise more.

From a TCM perspective the protocol was designed with the assumption that the knee pain was the primary problem. This may not have been the case for all participants. There is debate concerning whether the 36 questions of the RAND should be combined to form two aggregate scales, mental and physical health (Laucis et al., 2015). Part of this discussion is whether there is orthogonal or oblique relationship between these concepts. In TBSA mental and physical health are related. The physical knee pain can have an effect on the mental health, however, one must assess the locus of the problem. Individuals may have poor mental health primarily because of the knee pain. Alternatively they may have poor mental health due to other reasons, the knee pain exacerbates these problems - and poor mental health exacerbates the knee pain. The three participants whose MHC score was consistently low may have fit the latter scenario.

A future protocol could include the option to allow participants to have acupuncture with no protocol restrictions. The acupuncturist, who must remain blinded, could identify participants at the midpoint who they felt would benefit from a very different acupuncture approach. These participants would form a third arm to the study. They would serve as their own controls, by comparing semi-flexible to unrestricted acupuncture. Ultimately, the purpose of this study

was to find out whether moxibustion provides *additional* benefits. Patients for whom the semi-flexible point formula is ineffective should be excluded from the analysis.

VIII.4.1.Conclusion

The study has shown that the administrative procedures, the enrolment process and information supplied to participants were satisfactory and do not need to be adjusted. There were slight problems with the outcome measures. Ambiguity of questions 5 and 7 on the WOMAC questionnaire may lead some participants to answer 'different' questions at the various time points. The RAND-36 (v1.0) has eight sub-scales. Some of these were prone to ceiling and basement effects, and therefore may not be appropriate for measuring change within a clinical trial. The physical and mental health composites were not affected by this problem. Participants did not find completing the questionnaires a burden, and they appeared to be straightforward and easy to use. In addition both WOMAC and RAND-36 (v1.0) have been widely used and population based normative data are available. Therefore, both of these outcome measures would be suitable for a future RCT. It maybe possible to collect more information, potentially, participants could be asked keep a diary of medication use and activity.

The blinding procedures could be improved through better training and monitoring of the application of moxibustion. Blinding success may also be enhanced by informing participants that moxibustion may or may not be used at each appointment. A semi-flexible protocol may not be appropriate for all participants. A potential solution to this problem would be a third arm in the study design. Participants who did not respond initially could be reassigned to unrestricted acupuncture group. Prior to running a full scale trial a feasibility study should be conducted. This would provide information on the proportion of participants likely to allocated to the unrestricted acupuncture arm, and a power calculation. It would enable the enhanced blinding procedures to be evaluated.

IX.

Discussion

讨论

IX.1. Introduction

This chapter summarises the project's findings and original contribution to knowledge. Within the chapter the author refers back to the original aims and objectives, and identifies areas for future research

The project has developed new knowledge in three domains of acupuncture research; theory, practice, and clinical research. The methodological challenges of undertaking acupuncture research were discussed in [chapter-I](#). Potential solutions to these challenges were proposed and a new research framework established in [chapter-II](#). The research framework guided the design of subsequent phases of this project. The theoretical models, of the main styles of acupuncture, were examined in [chapter-III](#). The expert interviews and practitioner survey advanced understanding of acupuncture in practice ([chapters-V/VI](#)). The systematic review ([chapter-IV](#)) and the pilot study ([chapter-VIII](#)) developed new knowledge in the domain of clinical research. The translational analysis ([chapter-VII](#)) synthesises aspects of all three domains to produce new insights into the existing evidence-base. The translational analysis also shed light on potential shifts in theory and practice.

IX.2. Phase 1: Methodological challenges and solutions of acupuncture research

In phase-1 the literature on difficulties and controversies associated with acupuncture research were reviewed. [Chapter-I](#) presents a fresh analysis of the apparent incompatibility of TEAM and clinical research methods. Two factors were identified as presenting particular challenges. The first was a group of inter-related issues that derived from the characteristics of TEAM. The second was the more specific difficulty of developing a placebo control procedure. It was argued that within the context of acupuncture research ‘placebo’ is a meaningless concept.

A novel research framework was developed to accommodate these challenges and guide the subsequent phases of the project ([chapter-II](#)). In line with recommendations in the literature, the framework advocates a non-hierarchical circular relationship between methods and a translational research approach. The framework emphasises that component analysis should underpin acupuncture research and proposes new methods and axioms ([box-II.3](#)).

Within the framework a new schematic representation of the components of acupuncture is developed ([section-II.4.1](#)). The schema separates physically mediated from psychologically mediated components to better address the controversial issue of ‘placebo’. An axiom of the framework is that there is no single theory of TEAM or TBSA that can be falsified by means of scientific investigation ([section-II.2.1.b](#)). Instead, TBSA is identified as an area of research and an integral part of a translational approach ([section-II.4.6](#)). The concept of different levels of theories is introduced as a means of developing an overall theory of acupuncture ([section-II.4.2](#)). The TBSA diagnostic method - pattern differentiation - is conceptualised of as a staged process to facilitate the design and interpretation of research ([section-II.4.5](#)).

IX.3. Phase 2: Literature based research

The project has two literature based components, a narrative review of the theoretical approaches to treating osteoarthritis of the knee ([chapter-III](#)) and a systematic review ([chapter-IV](#)). These are collectively described as phase-2 of the project. The primary objective of phase 2 was to identify a suitable condition for the pilot study, suitability was assessed on a number of criteria ([section-II.5.2](#)). The secondary objective was to gather information for the translational analysis presented in [chapter-VII](#).

The criteria to assess the suitability of a condition for the pilot study evolved from the framework which proposes: component efficacy trials investigating the physical variables of acupuncture should, ideally, select a bing-illness for which a semi-flexible protocol would have a high degree of external validity ([section-II.4.5.b](#)). A core theory within the TCM style, described as the *standard theory* within this project, was identified. The narrative review suggested that using the standard theory, osteoarthritis of the knee was a suitable bing-illness for the RCT ([section-III.7.2](#)).

There were, however, caveats. The review encompassed other styles of practice, including other traditionally based systems of acupuncture (TBSA) (Five-elements, Stems and branches, Japanese) as well as medical acupuncture. The non-TCM TBSAs appear to place greater emphasis on treating the ben-root. A semi-flexible protocol would not be generalisable to these styles of practice. Similarly, it was also not clear how important treating ben-root was among TCM acupuncturists in practice. The key limitation of the review was that it could only deal with theory, it could not provide any information regarding acupuncture in actual practice. Hence, the need to conduct the primary research consisting of expert interviews and a practitioner survey was established.

The framework proposes that component analysis should be central to acupuncture research. The clinical reasoning process of the *standard theory* was

used to identify a method of categorising acupuncture points; the *modes of acupuncture point selection* (section-III.5.5.a). These modes are *components* of the clinical reasoning processing. The *modes of acupuncture point selection* would be recognisable to a trained acupuncturist, but explicitly incorporating them into research is a novel development.

A weakness of the narrative review is that it is TCM centric. If a researcher more familiar with a non-TCM style had undertaken the review, then potentially the modes may have been different. TCM is the dominant style of practice and, consequently, there are obvious benefits in utilising TCM theory as the basis for analysing the components of clinical reasoning. Nevertheless, it has meant that within this project non-TCM styles have been analysed in comparison to TCM, rather than as independent theoretical models.

The existing evidence regarding the hypothesis of the prospective RCT (box-II.4) was evaluated by conducting a systematic review. As far as the researcher is aware, this was the first such review in English. A total of 34 studies were reviewed, which collectively had 3134 participants. Serious methodological and reporting shortcomings were found in these studies (section-IV.4). In addition, none of them had used blinding procedures similar to those tested in the pilot study. The need to design and conduct a controlled trial to investigate warm needle acupuncture for osteoarthritis of the knee was thereby established. The systematic review uncovered a difficulty in using Chinese language databases (section-IV.6). Awareness of this issue will be of benefit to researchers in the future.

Two other criteria were assessed to ensure osteoarthritis of the knee was an appropriate condition for the pilot study. Firstly, the disease burden and available treatments ought to warrant the investigation of a novel therapy, this was confirmed in section-III.3. Secondly, comparative effectiveness trials should show that the therapy in its entirety has clinically relevant benefits prior to

component efficacy studies (Craig et al., 2013). This was confirmed in [section-IV](#).

2.

Box-IX.1: Summary of Phase 2: Literature based research

Primary objective: Osteoarthritis of the knee was suitable for the RCT

- ▶ A semi-flexible protocol based on standard theory (TCM) should have external validity ([box-III.4](#))
- ▶ The systematic review demonstrated that the existing evidence was insufficient to draw any conclusions regarding the hypothesis: warm needle acupuncture leads to greater clinical benefits than needle acupuncture
- ▶ The disease burden and current treatment options warrant investigation of a novel therapy
- ▶ Comparative effectiveness trials indicated that acupuncture has clinically relevant benefits for osteoarthritis of the knee

Knowledge development

- ▶ Incorporation of the *modes of acupuncture point selection* into the research framework
- ▶ Identification of specific challenges when using Chinese language databases for systematic reviews
- ▶ Identification of poor reporting and methodological standards of Chinese research - in particular the outcome measures

IX.4. Phase 3: Primary research - acupuncture in practice

The main objective of phase-3, expert interviews and practitioner survey, was to develop an understanding of acupuncture in practice. The two pieces of work are closely related. The expert interviews, in conjunction with the narrative review of acupuncture theory ([chapter-III](#)), was used to develop the survey questionnaire.

IX.4.1. Overview of new knowledge from the practitioner survey

Two hundred and eleven completed questionnaires were returned in the practitioner survey, which makes it possibly the largest survey of its kind. Whilst there were differences between the three groups (medical, TCM, and combined non-TCM acupuncture), within each group there was considerable heterogeneity. This heterogeneity means that styles of practice can not be clearly delineated when considering the needling components of acupuncture. The use or otherwise of non-needle techniques, such as moxibustion and cupping, differentiate TBSA from medical acupuncture. The survey provided a wide range of new knowledge about acupuncture in practice for osteoarthritis of the knee. This enabled benchmarks to be established that indicate how commonly each component of acupuncture is used in practice. Knowledge developed in phase-3 suggests that there are no discrete styles of practice but rather a spectrum. At one end of the spectrum practitioners only use local points, at the other end distal points will nearly always be selected and local points occasionally omitted. The poles of the spectrum can be considered as either reductionist-holistic, medical - traditional, or simple - complex.

IX.4.2. Psychological factors influence the physical treatment

The expert interviews provided an enhanced understanding of the clinical decision-making of acupuncturists. Information derived from the interviews suggests that acupuncturists may adapt the physical treatment based on a number of psychological factors ([section-V.5.4](#)). This led to the conceptualisation

of additional *modes of acupuncture point selection*. These include adapting treatment because the patient is: new to acupuncture, anxious, sensitive to needling. The physical treatment may also be altered when a practitioner believes the patient would consider acupuncture points distal to the knee to be disconcerting, as they are perceived as not logical.

The influence of psychological factors on the physical process of acupuncture is an important theme that emerged from phase-3. The practitioner survey demonstrates that reduction in the number of needles used is common in the early stages of treatment. It also shows that acupuncturists may adjust the physical variables, point selection and degree of stimulation, in response to a patient's anxiety and sensitivity ([section-V.5.4](#)). For example, the addition of specific acupuncture points thought to help reduce anxiety; or a reduction in the number of needles inserted. Reducing the strength of stimulation for sensitive patients and a minimal approach during the first few sessions has been described elsewhere (MacPherson et al., 2006).

The project focuses on the physical components of acupuncture ([section-II.2.2](#)). As a result there are no qualitative in-depth investigations into the beliefs and clinical reasoning of practitioners. This is a limitation of the project. Future research could explore the beliefs of acupuncturists regarding where they stand on the spectrum and the factors that have influenced their thinking. Some of the factors that may influence clinical reasoning are: a biomedical understanding of the body; clinical experience; time pressures. It seems plausible that the beliefs of individual acupuncturists evolve over time. This may explain why medical acupuncturists appear to have adopted various practices from TBSA that would not necessarily be anticipated from the narrative review ([chapter-III](#)). And why some TBSA practitioners appear to predominantly deliver a reductionist style of acupuncture ([chapter-VI](#)).

IX.4.3. The symbolic nature of point selection

The potential for patients' beliefs to influence the physical process is an interesting finding (section-V.5.4.d). The selection of acupuncture points close to the knee or at distal locations have symbolic implications. If the acupuncturist believes the patient would find the symbolism of distal points challenging, they may adapt the physical treatment accordingly. The survey results suggest that about a third practitioners occasionally adjust treatment because the patient might perceive the treatment as illogical (section-VI.4.7.a). This was, perhaps, a difficult concept to include within a predominantly quantitative survey and qualitative research is necessary to further explore the clinical reasoning of practitioners.

Along with many other CAM therapies acupuncture can be described as 'holistic': a therapy that does not focus on the part of the body that appears to be 'ill', but rather attempts to view the body as a whole. The holistic approach and the contrast with the perception of biomedicine as being reductionist, focused on parts not the whole, can be one of the factors that leads an individual to seek acupuncture for their condition (Jakes et al., 2014). The use of distal points symbolises a more holistic approach, whereas local points symbolises a reductionist understanding. Some patients may find notions of holism difficult to accept (Jakes et al., 2014). Patients may or may not align themselves with the metaphysical/ spiritual dimension of TBSA.

The reasons patients choose specific acupuncturists are yet to be fully explored, although it appears the reassurance of a medically trained practitioner influences some (Bishop et al., 2011b). Patients vary in their feelings towards acupuncture, some perceive acupuncture as a good fit for themselves whilst others are concerned about exotic or mystical health theories (Bishop and Lewith, 2013). As patients go through the process of choosing an acupuncturist, speaking to friends and relatives, reading leaflets and websites, it seems reasonable to assume they will choose a practitioner they perceive as having a style of practice that accords with their meta-physical view point. In other

words, those patients who find the exotic-mystical nature of acupuncture a pull factor will seek an acupuncturist who presents themselves as such; and may receive a treatment with a wide distribution of points. Those patients who are concerned by the exotic-mystical nature will seek out a medical acupuncturist; and may receive a treatment with a narrow distribution of points located around the knee. Consequently if an individual decides to try acupuncture, the physical distribution of points used may depend more on their meta-physical view point than their clinical signs and symptoms.

Patients views on acupuncture are not static. At the first appointment they do not necessarily know anything about the theories or potential mechanisms of acupuncture; but there can be a educative process (Jakes et al., 2014). One of the findings of the experts interviews also suggests that there was educative process ([section-V.5.4.d](#)). As the patient becomes more comfortable with the language and a holistic approach in general, the physical treatment may change.

IX.4.4. Physiological responses to the physical process may be moderated by psychological factors

The motivations of acupuncturists in making adjustments due to psychological factors ([section-V.5.4](#)) may be simply that they do not wish to discourage patients from continuing the course of treatment. However, it is also possible that these modifications actually improve the treatment. There is evidence to suggest that high stimulation intensity coupled with an active coping strategy and low stimulation intensity coupled with a passive coping strategy should be recommended for improved therapeutic effects (Lee et al., 2014). As the patient develops a relationship of trust with the acupuncturist this may act as an active coping strategy. The patient has confidence that the acupuncturist will respond to requests to reduce the intensity of stimulation, if needed. In the early stages of a course of treatment this confidence has not been established, therefore, a lower intensity of stimulation should be administered. In other words, if a patient feels in control then a stronger physical stimulation can be administered.

The selection of points distal to the knee may increase feelings of anxiety in an individual who finds it difficult to believe these points may be useful. This anxiety could be potentially be associated with feelings of 'lack of control'. This in turn may influence the physiological effects of the needle insertion. Conversely, an individual who is drawn to the holistic aspects of acupuncture may not only be comfortable with the use of distal points but is reassured, as this meets their expectations. It has been suggested that the psychological effects of taking pills will be different for an individual living in a society where television commercials attest to the effectiveness of medication when compared to those of an individual who lives in society where evil spirits are widely accepted as being the cause of illness (Frenkel, 2008). Acupuncturists may intuitively, or consciously, moderate the physical variables of treatment to account for the meta-physical beliefs of their patients. Ironically, for some patients a treatment that eschews the TBSA individualised selection of points, in favour of just points around the knee, may be more holistic. As the simple - local treatment takes account of the patient's meta-physical beliefs. Further research is required to understand how the nexus of psychological factors of belief, anxiety, sensitivity and sense of control impact on the physiological response to the needling process.

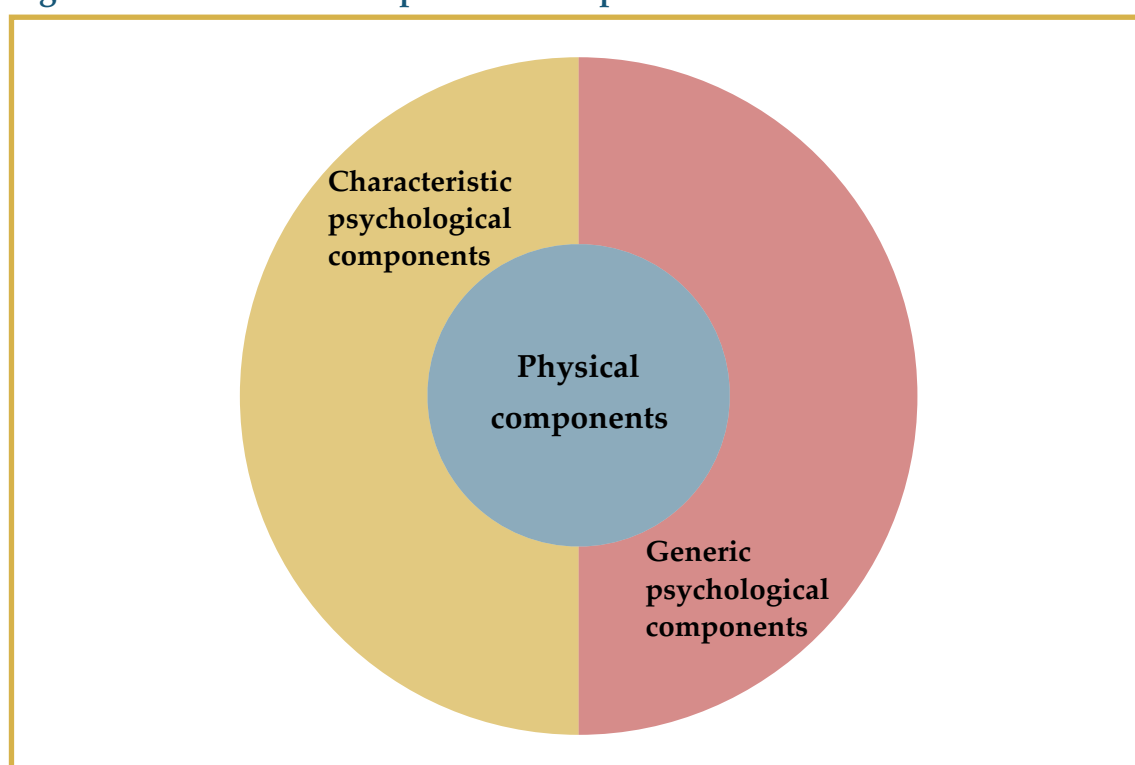
Individualised care is usually seen as a core approach within TBSA, and is derived from a detailed case history and the traditional diagnostic process (MacPherson et al., 2006). The individualised approach is one of the valued attributes of acupuncture in the private sector, along with empowerment and shared decision making (Bishop et al., 2011a). The moderation of treatment to account for sensitivity and anxiety, as well as moderation to align with the patient's meta-physical beliefs, can also be seen as aspects of individualised care. Possibly, individualisation contributes to the clinical benefits of acupuncture mediated *directly* through the psychological effects of empowerment and shared decision making. However, these psychological

factors may *moderate the physiological response* to the physical components, such as needle insertion.

IX.4.5. Schema of acupuncture components - revised

The clear division of psychological and physical components was an innovation of the research framework (box-II.3). Interestingly, it was this very division that facilitated the emergence of a new understanding of the clinical decision-making process. Analysing acupuncture in terms of purely physical components highlighted that these variables were dependent on the assessment of psychological factors. Moreover, the physiological response to the physical actions (e.g. needle insertion) may be moderated by the psychological state of the patient. The framework must now be updated to account for the highly complex relationship between the psychological and physical components. Figure-IX.1 illustrates a revised version of the schema originally presented in figure-II.6. The physical components have been moved to the centre to highlight that they can only exist within the context psychological factors.

Figure-IX.1: Schema of acupuncture components - revised version



Component efficacy trials can investigate the *specific effects* of individual components. If psychological components, both characteristic and generic, are held in equal between two groups, then any demonstrated efficacy must be due to the physical components. These could be described as placebo-controlled trials (section-I.4.4). Ideally the experimental and control conditions should differ only in respect to one physical component such as insertion/ non-insertion. The efficacy of psychological components can be investigated by holding equal the physical components between groups. In other words, use the same acupuncture protocol in both experiment and control groups whilst manipulating one of the psychological factors - component X. However, interpretation of results is potentially complex. Any demonstrated efficacy might be due to the *direct psychological benefit* of component X, or component X might *moderate the physiological response* to the physical acupuncture.

Box-IX.2: Summary of Phase 3: Acupuncture in practice

Knowledge development

- ▶ Additional *modes of acupuncture points selection/adaptation* were identified that enhanced the understanding of the clinical reasoning process. These were psychological factors that had not been identified in the narrative review
- ▶ There is a high degree of intra-style heterogeneity
- ▶ When considering the needling process, styles of practice cannot be clearly delineated
- ▶ Acupuncture styles should be conceived as a spectrum rather than discreet styles of practice
- ▶ TBSA can be delineated from medical acupuncture by the use of moxibustion, cupping and heat lamps

IX.5. Phase 4: Translational analysis and pilot study

IX.5.1. Translational analysis

Clinical trials can only be effectively evaluated with a contextual understanding of theory and practice. There are two key issues: the risk of bias from sham-acupuncture procedures; the external validity of the real acupuncture ([section-II.4.6.b](#)). A limitation of the systematic reviews that have evaluated acupuncture for osteoarthritis of the knee is they did not assess the external validity of the acupuncture or the risk of bias from the sham procedures ([section-IV.2](#)).

A list of 22 criteria to assess both sham and real acupuncture was developed by synthesising knowledge from phases 2 and 3 of this project. These criteria were then used to evaluate the procedures used in high-quality clinical trials. Each criterion was assessed as being high, unclear, and low risk. Benchmarks were obtained from the practitioner survey that informed the analysis.

The translational analysis demonstrated that acupuncture protocols used in clinical trials have differing degrees of external validity. Similarly, the sham procedures appear to be associated with different levels of risk of bias. This suggests that previous interpretations tend to underestimate the efficacy of acupuncture. The analysis identified clinical trials that combined poor external validity with a high risk bias from the sham procedures, such trials should potentially be excluded from systematic reviews.

The translational analysis not only developed insights regarding the evidence-base for osteoarthritis of the knee, but also is a proof of concept: that demonstrates a new method of assessing acupuncture and sham procedures. This method could be applied to evaluate systematic reviews for other conditions, as well as support the development of clinical trials. Previous evaluation methods used generic criteria, that are applicable for any condition. By developing criteria that are specific to osteoarthritis of the knee a more nuanced analysis was possible. The 22 criteria could potentially be used to

assess other painful conditions, however, the benchmarks derived from the survey would no longer be applicable. Ideally, new criteria should be developed for each condition using literature reviews of theory, interviews and surveys. Criteria to assess non-painful conditions may differ considerably from the 22 that evaluated acupuncture for osteoarthritis.

The degree to which clinical research is generalisable was also evaluated. It has been recommended that the style of practice should be clearly reported (MacPherson et al., 2010). The majority of trials reported that their acupuncture protocols were based on TCM. However, some protocols focus on the affected limb (Suarez-Almazor et al., 2010) whilst others use distal points to address the ben-root (Scharf et al., 2006). The analysis highlighted that these trials tested different kinds of acupuncture, and the simple description of TCM acupuncture masks heterogeneity within the systematic review by Vickers et al (2017). The Vickers et al (2017) review can be generalised to medical acupuncture and to some extent TCM, however, it does not appear to reflect other TBSA. Because of the heterogeneity within styles and the difficulty in delineating styles ([section-VI.5](#)), statements on generalisability must be interpreted with caution.

The differences between acupuncture in practice and the recommendations set out in textbooks were also examined. This comparison of theory and practice identified a number of areas for future research, that would usefully be explored using qualitative methods. Research questions emerge wherever practice appeared to differ from the recommended theory. For example, why are back-shu used less frequently than the theory would suggest? How do medical practitioners incorporate ideas such as 'tonification' into the biomedical model?

Box-IX.3: Summary of the translational analysis**Knowledge development**

- ▶ A novel method of evaluating external validity and generalisability, as well as the risk of bias of sham acupuncture was developed and tested
- ▶ Acupuncture procedures in clinical trials were shown to have different degrees of external validity
- ▶ Minimal acupuncture sham procedures were shown to have different levels of risk of bias
- ▶ Clinical research reflects medical acupuncture as much as, if not more than, TBSAs
- ▶ A group of acupuncture points known as back-shu appear to be used less frequently than textbook recommendations would suggest

IX.5.2. The pilot study

The design of the pilot study was inspired by the research framework developed in the first two chapters. Clinical trials with the capacity to demonstrate that the physical components of acupuncture have specific effects are particularly important ([section-II.2.2](#)). They are able to show that acupuncture is not ‘just a placebo’, because the spectrum of factors that constitute the placebo effect are held equal between groups. Consequently, such trials should be more influential in persuading those who are currently sceptical about acupuncture ([section-II.2.1](#)). In addition, they provide useful information for practitioners regarding best practice and can be linked with mechanism studies ([section-II.2.2](#)). The protocol designed and tested within this project is one such study. It tests the *specific effect* of heat conducted along the needle into acupuncture points local to the knee, in patients with osteoarthritis. It is the first study of its kind, with novel methods of blinding acupuncturist and participants.

The pilot study protocol was successfully evaluated. Many of the administrative processes were proven to be sound. Potential improvements to the protocol were identified. In particular, enhancement of the blinding procedures and the reallocation of non-responders into a third arm.

In the study, acupuncture point selection was limited to local points. Naturally this meant that other modes of acupuncture point selection, meridian, general health and other signs and symptoms, were not used. These limitations were put in place as a compromise to facilitate the blinding procedure and enhance replicability. Osteoarthritis of the knee was chosen on the belief that the exclusion of these modes of point selection would have minimal impact on external validity ([section-III.7.2](#)). The veracity of this assumption was initially judged on the basis of the theory. The practitioner survey demonstrated a protocol of local points only would lack external validity for some practitioners, but not all. In addition Chinese clinical trials almost exclusively used local points ([section-VII.4.4](#)). Therefore, the assumption that an acupuncture protocol limited to local points would have external validity appears to be sound.

It is important not to equate external validity with optimal practice. Ultimately the clinical impact of the absence of other modes of acupuncture point selection, such as meridian points, is unknown. Component efficacy trials are needed to compare the use of local points alone to local and distal combined. Some research has been conducted but at present it is not possible to draw any conclusions (Wong Lit Wan et al., 2015).

The pilot study highlighted ethical issues that need to be considered when designing component efficacy trials. Due to the heterogeneity of practice it is difficult to know which ‘missing’ elements participants should be informed about. Should, for example, participants have been informed that not using distal points on the affected meridian might reduce the overall effectiveness? If so, then perhaps they should have also been informed that the protocol lacked: back-shu points; points at the elbow; ear seeds; points relating to their date of birth. The list of components that might be restricted/changed is very long. We also have to consider that the addition of an element might in some cases be a less effective treatment. For example, using points to treat the underlying condition may not be appropriate for some participants, if they are sensitive to

needles. Ultimately, we have returned to one of the challenges discussed in chapter-I: what is acupuncture? ([section-I.2.1.b](#)). It would be unreasonable to expect participants to grapple with all the disputes within the acupuncture profession about what constitutes best practice. In the pilot study, participants were informed that they would ‘have a course of acupuncture’. On reflection, this does seem justified on the basis that some practitioners, not least those conducting more pragmatic clinical research in China, simply use local points. Practitioners who feel that the use of distal points is particularly important for osteoarthritis of the knee, naturally, may take a different view. The description of the interventions will be an important ethical issue to consider in future component efficacy trials.

The pilot study also identified another ethical issue. Whilst the inclusion of meridian points may have lead to a more effective treatment for those who responded, for a few the protocol appeared to be ineffective. The participants who had persistently low MHC scores may well have been depressed. In terms of TBSA this would indicate that the bing-illness should be reassessed. In other words, the bing-illness would be depression rather than knee pain. This would necessitate a radically different prescription of acupuncture points. It may be difficult to identify at baseline participants for whom the protocol is inappropriate. Some participants who had low MHC scores at baseline improved quickly with treatment. Researchers will need to consider whether it is ethical to continue treatment if the semi-flexible protocol does not appear to be working. The option to allow participants to be reallocated to an unrestricted acupuncture arm offers a method of dealing with this problem. As well as providing a solution for the ethical problem, the addition of a third arm would enable comparisons to be made between restricted and unrestricted acupuncture protocols. Future clinical trials that use fixed or semi-flexible acupuncture protocols should consider including a third arm for those who do not respond to the initial sessions.

It is possible that the same phenomenon occurred in previous studies of acupuncture for osteoarthritis of knee, or other painful conditions. In this study 3/14 participants had consistently low MHC scores throughout. By chance these were all in the warm needle acupuncture group. In larger trials randomisation should ensure that potentially depressed participants are evenly spread between groups. Therefore, this issue should have minimal impact on that between groups comparisons. However, it does mean the overall effectiveness of acupuncture may be underestimated. This would certainly be the case if these participants respond better once re-allocated to a unrestricted acupuncture arm.

Box-IX.4: Summary of the pilot study

Knowledge development

- The general administrative procedures were sufficient and do not need to be adjusted
- The blinding procedures were partially successful, areas for improvement were identified
- A third arm should be included to account for participants who do not respond to the semi-flexible protocol
- Inter-rater reliability of the pattern differentiation should be assessed

Limitations

- Insufficient numbers were recruited to enable a power calculation

IX.5.3. Final Summary

The original objective of this project was to design a randomised-controlled trial protocol to investigate the clinical benefits of the additional use of moxibustion. The protocol was designed and tested. Problematic issues were identified, and potential solutions to these were proposed. The protocol could be used for a prospective larger scale RCT.

The methodological challenges of acupuncture research were debated and a new research framework developed. The framework analysed the acupuncture in terms of its components and was used to guide the subsequent phases of the project, including the RCT pilot study. Knowledge of practice, developed from two primary research studies, was combined with TBSA theory to re-evaluate the existing evidence base for osteoarthritis of the knee. This analysis demonstrated that acupuncture protocols have different degrees of external validity. In addition, the sham procedures appear to have different levels of risk of bias. This suggests the current practice of interpreting research as a simple dichotomy of acupuncture versus sham-placebo underestimates the efficacy of acupuncture.

The research framework designed for this project originally clearly separated physically and psychologically mediated components. This was found to have limitations, therefore, the framework was revised to show that the physical can only exist *within* the psychological components. The framework and the methods used within this study could be applied to acupuncture for other conditions.

References

- AACP. (2014) *Acupuncture Association of Chartered Physiotherapists*. Available from: <http://www.aacp.org.uk> [Accessed 7 December 2015]
- Abhishek, A. & Doherty, M. (2013) Mechanisms of the placebo response in pain in osteoarthritis, *Pain in Osteoarthritis*, 21, pp. 1229-1235. <http://dx.doi.org/10.1016/j.joca.2013.04.018>
- Acupuncture Academy. (2018) *Acupuncture Course | The Acupuncture Academy*. Available from: <http://theacupunctureacademy.org.uk/online-acupuncture-course/> [Accessed 23 February 2018]
- Aird, M., Cobbin, D. M. & Rogers, C. (2002) A study of the relative precision of acupoint location methods, *Journal of alternative and complementary medicine*, 8, pp. 635-642.
- Aird, M., Coyle, M., Cobbin, D. M. & Zaslowski, C. (2000) A Study of the Comparative Accuracy of Two Methods of Locating Acupuncture Points, *Acupuncture In Medicine: Journal Of The British Medical Acupuncture Society*, 18, pp. 15-21.
- Alper, B. S., Shah, A., Malone-Moses, M., Manheimer, E. W. & Ehrlich, A. (2016) Point-of-care application of: Guidelines and evidence on acupuncture for chronic low back pain, *European Journal of Integrative Medicine*, 8, pp. 326-328. <http://dx.doi.org/10.1016/j.eujim.2016.07.029>
- Alraek, T. (2014) Designing clinical studies that take into account traditional East Asian medicine's systems and methods — with focus on pattern identification, *Chinese Journal of Integrative Medicine*, 20, pp. 332-335. Doi: 10.1007/s11655-014-1807-5.
- Altman, D. G. & Bland, J. M. (1995) Absence of evidence is not evidence of absence, *BMJ Clinical Research*, 311, pp. 485-485. Doi: 10.1136/bmj.311.7003.485.
- Altman, R., Asch, E., Bloch, D., Bole, G., Borenstein, D., Brandt, K., Christy, W., Cooke, T. D., Greenwald, R., Hochberg, M. & et al. (1986) Development of criteria for the classification and reporting of osteoarthritis. Classification of osteoarthritis of the knee. Diagnostic and Therapeutic Criteria Committee of the American Rheumatism Association, *Arthritis and Rheumatism*, 29, pp. 1039-49. Doi: 10.1002/art.1780290816.
- American College of Rheumatology. (1986) *Classification Criteria for Rheumatic Diseases: Idiopathic OA of the knee*. Available from: <https://www.rheumatology.org/Practice-Quality/Clinical-Support/Criteria/ACR-Endorsed-Criteria> [Accessed 5 January 2014].
- Appleyard, I., Crichton, N. & Robinson, N. (2016) Warm needle acupuncture vs. needle acupuncture for osteoarthritis of the knee: A pilot study protocol, *European Journal of Integrative Medicine*, 8, pp. 407-413. <http://dx.doi.org/10.1016/j.eujim.2016.07.031>
- Appleyard, I., Lundeberg, T. & Robinson, N. (2014) Should systematic reviews assess the risk of bias from sham-placebo acupuncture control procedures?, *European Journal of Integrative Medicine*, 6, pp. 234-243. Doi: 10.1016/j.eujim.2014.03.004.
- Araim, M., Campbell, M. J., Cooper, C. L. & Lancaster, G. A. (2010) What is a pilot or feasibility study? A review of current practice and editorial policy, *BMC Medical Research Methodology*, 10, pp. 67-67. Doi: 10.1186/1471-2288-10-67.

- Asai, A. & Kadooka, Y. (2013) Re-examination of the ethics of placebo use in clinical practice, *Bioethics*, 27, pp. 186-193. Doi: 10.1111/j.1467-8519.2011.01943.x.
- BAAB. (No date): British Acupuncture Accreditation Board. Available from: <https://www.baab.co.uk/> [Accessed 20 May 2015]
- Baldry, P. E. (2004) *Acupuncture, Trigger Points and Musculoskeletal Pain*. (Acupuncture, Trigger Points, & Musculoskeletal Pain). 3 ed. Churchill Livingstone.
- Bang, H., Flaherty, S. P., Kolahi, J. & Park, J. (2010) Blinding assessment in clinical trials: A review of statistical methods and a proposal of blinding assessment protocol, *Clinical Research & Regulatory Affairs*, 27, pp. 42-51. Doi: 10.3109/10601331003777444.
- Bellamy, N. (2015) *WOMAC® Osteoarthritis Index - User Guide*, Queensland, Australia: Professor Nicholas Bellamy.
- Bellamy, N. (2016) *WOMAC Knee and Hip Osteoarthritis Index*. Available from: <http://www.womac.org/> [Accessed 15 September 2017]
- Bellamy, N., Wilson, C. & Hendrikz, J. (2011) Population-based normative values for the Western Ontario and McMaster (WOMAC) Osteoarthritis Index: part I, *Seminars in arthritis and rheumatism*, 41, pp. 139-148. Doi: 10.1016/j.semarthrit.2011.03.002.
- Benedetti, F., Arduino, C. & Amanzio, M. (1999) Somatotopic activation of opioid systems by target-directed expectations of analgesia, *The Journal Of Neuroscience: The Official Journal Of The Society For Neuroscience*, 19, pp. 3639-3648.
- Berman, B. M., Lao, L., Langenberg, P., Lee, W. L., Gilpin, A. M. K. & Hochberg, M. C. (2004) Effectiveness of acupuncture as adjunctive therapy in osteoarthritis of the knee: a randomized, controlled trial, *Annals of Internal Medicine*, 141, pp. 901-910.
- Billingham, S. A., Whitehead, A. L. & Julious, S. A. (2013) An audit of sample sizes for pilot and feasibility trials being undertaken in the United Kingdom registered in the United Kingdom Clinical Research Network database, *BMC Medical Research Methodology*, 13, pp. 104. Doi: 10.1186/1471-2288-13-104.
- Bilton, K. & Zaslowski, C. (2016) Reliability of Manual Pulse Diagnosis Methods in Traditional East Asian Medicine: A Systematic Narrative Literature Review, *Journal of alternative and complementary medicine*, 22, pp. 599-609. Doi: 10.1089/acm.2016.0056.
- Birch, S. (1997) Testing the claims of traditionally based acupuncture, *Complementary therapies in medicine*, 5, pp. 147-151. Doi: 10.1016/S0965-2299(97)80057-X.
- Birch, S. (2004) Clinical research on acupuncture. Part 2. Controlled clinical trials, an overview of their methods, *Journal of alternative and complementary medicine*, 10, pp. 481-498. <https://doi.org/10.1089/1075553041323911>
- Birch, S. (2006) A review and analysis of placebo treatments, placebo effects, and placebo controls in trials of medical procedures when sham is not inert, *Journal of alternative and complementary medicine*, 12, pp. 303-310.
- Birch, S. (2009) Filling the whole in acupuncture: Part 1:2 What are we doing in the supplementation needle technique? Scientific perspectives, *European Journal of Oriental Medicine*, 6, pp. 18-27.
- Birch, S. (2015a) Historical and clinical perspectives on de Qi: Exposing limitations in the scientific study of de Qi, *Journal of alternative and complementary medicine*, 21, pp. 1-7. Doi: 10.1089/acm.2014.0267.

- Birch, S. (2015b) Understanding Qi in clinical practice - Perspectives from an acupuncture scholar-practitioner, *Journal of Chinese Medicine*, pp. 62-67.
- Birch, S. & Alraek, T. (2014) Traditional East Asian medicine: how to understand and approach diagnostic findings and patterns in a modern scientific framework?, *Chinese Journal of Integrative Medicine*, 20, pp. 336-40. Doi: 10.1007/s11655-014-1809-3.
- Birch, S., Alraek, T. & Lee, M. S. (2016) Challenges for clinical practice guidelines in traditional medicines: The example of acupuncture, *European Journal of Integrative Medicine*, 8, pp. 332-336. <http://dx.doi.org/10.1016/j.eujim.2016.07.032>
- Birch, S., Hesselink, J. K., Jonkman, F. A. M., Hekker, T. A. M. & Bos, A. (2004) Clinical research on acupuncture. Part 1. What have reviews of the efficacy and safety of acupuncture told us so far?, *Journal of alternative and complementary medicine*, 10, pp. 468-480.
- Birch, S. & Ida, J. (1998) *Japanese Acupuncture: A Clinical Guide*. Brookline, Massachusetts: Paradigm Publications.
- Birch, S. & Sherman, K. (1999) Zhong Yi acupuncture and low-back pain: traditional Chinese medical acupuncture differential diagnoses and treatments for chronic lumbar pain, *Journal of alternative and complementary medicine*, 5, pp. 415-425.
- Birch, S. J. & Felt, R. L. (1999) *Understanding acupuncture*. Edinburgh; New York: Churchill Livingstone.
- Birkeflet, O., Laake, P. & Vøllestad, N. (2011) Low inter-rater reliability in traditional Chinese medicine for female infertility, *Acupuncture In Medicine: Journal Of The British Medical Acupuncture Society*, 29, pp. 51-57. Doi: 10.1136/aim.2010.003186.
- Birkeflet, O., Laake, P. & Vøllestad, N. K. (2014) Poor multi-rater reliability in TCM pattern diagnoses and variation in the use of symptoms to obtain a diagnosis, *Acupuncture In Medicine: Journal Of The British Medical Acupuncture Society*, 32, pp. 325-332. Doi: 10.1136/acupmed-2013-010473.
- Bishop, F. L., Barlow, F., Coghlan, B., Lee, P. & Lewith, G. T. (2011a) Patients as healthcare consumers in the public and private sectors: a qualitative study of acupuncture in the UK, *BMC Health Services Research*, 11, pp. 129-129. Doi: 10.1186/1472-6963-11-129.
- Bishop, F. L., Jacobson, E. E., Shaw, J. & Kaptchuk, T. J. (2012a) Participants' experiences of being debriefed to placebo allocation in a clinical trial, *Qualitative Health Research*, 22, pp. 1138-1149. Doi: 10.1177/1049732312448544.
- Bishop, F. L., Jacobson, E. E., Shaw, J. R. & Kaptchuk, T. J. (2012b) Scientific tools, fake treatments, or triggers for psychological healing: How clinical trial participants conceptualise placebos, *Social science & medicine*, 74, pp. 767-774. <http://dx.doi.org/10.1016/j.socscimed.2011.11.020>
- Bishop, F. L. & Lewith, G. T. (2008) A review of psychosocial predictors of treatment outcomes: what factors might determine the clinical success of acupuncture for pain?, *Journal Of Acupuncture And Meridian Studies*, 1, pp. 1-12. Doi: 10.1016/S2005-2901(09)60001-7.
- Bishop, F. L. & Lewith, G. T. (2013) Patients' preconceptions of acupuncture: a qualitative study exploring the decisions patients make when seeking acupuncture, *BMC Complementary And Alternative Medicine*, 13, pp. 102-102. Doi: 10.1186/1472-6882-13-102.

- Bishop, F. L., Massey, Y., Yardley, L. & Lewith, G. T. (2011b) How patients choose acupuncturists: a mixed-methods project, *Journal of alternative and complementary medicine*, 17, pp. 19-25. Doi: 10.1089/acm.2010.0061.
- Bishop, F. L., Zaman, S. & Lewith, G. T. (2011c) Acupuncture for low back pain: A survey of clinical practice in the UK, *Complementary therapies in medicine*, 19, pp. 144-148. Doi: 10.1016/j.ctim.2011.03.004.
- Bjoridal, J. M., Klovning, A., Ljunggren, A. E. & Slørdal, L. (2007) Short-term efficacy of pharmacotherapeutic interventions in osteoarthritic knee pain: A meta-analysis of randomised placebo-controlled trials, *European Journal of Pain*, 11, pp. 125-138. Doi: 10.1016/j.ejpain.2006.02.013.
- Bjoridal, J. M., Ljunggren, A. E., Klovning, A. & Slørdal, L. (2004) Non-steroidal anti-inflammatory drugs, including cyclo-oxygenase-2 inhibitors, in osteoarthritic knee pain: meta-analysis of randomised placebo controlled trials, *BMJ Clinical Research*, 329, pp. 1317-1317. Doi: 10.1136/bmj.38273.626655.63.
- Bostick, N. A., Sade, R., Levine, M. A. & Stewart, D. M., Jr. (2008) Placebo use in clinical practice: report of the American Medical Association Council on Ethical and Judicial Affairs, *The Journal Of Clinical Ethics*, 19, pp. 58-61.
- Bovey, M. (2016) Acupuncture for osteoarthritis in the UK: A turning point for NICE?, *European Journal of Integrative Medicine*, 8, pp. 337-341. <http://dx.doi.org/10.1016/j.eujim.2016.08.001>
- Bowling, A. (2009) Research methods in health: investigating health and health services. 3rd ed. England: McGraw-Hill Education.
- Bradnam, L. V. (2011) A biopsychosocial clinical reasoning model for Western acupuncture, *Physical Therapy Reviews*, 16, pp. 138-146. Doi: 10.1179/1743288X11Y.0000000011.
- Brinkhaus, B., Becker-Witt, C., Jena, S., Linde, K., Streng, A., Wagenpfeil, S., Irnich, D., Hummelsberger, J., Melchart, D. & Willich, S. N. (2003) Acupuncture Randomized Trials (ART) in patients with chronic low back pain and osteoarthritis of the knee – design and protocols, *Forsch Komplementarmed Klass Naturheilkd*, 10, pp. 185-191. Doi: 10.1159/000073474.
- British Acupuncture Council. (2011) *BACc Member Survey 2011*. Available from: <https://www.acupuncture.org.uk/category/members-section/members-about-us/members-about-us-membership-surveys.html> [Accessed 7 December 2015]
- British Acupuncture Council. (2016) *BACc Professional Codes*. Available from: <http://www.acupuncture.org.uk/public-content/effective-practice/bacc-professional-codes.html> [Accessed 14 September 2017].
- Buetow, S. (2010) Thematic analysis and its reconceptualization as 'saliency analysis', *Journal of Health Service Research and Policy*, 15, pp. 123-5. Doi: 10.1258/jhsrp.2009.009081.
- Cao, L., Zhang, X.-L., Gao, Y.-S. & Jiang, Y. (2012) Needle acupuncture for osteoarthritis of the knee. A systematic review and updated meta-analysis, *Saudi Medical Journal*, 33, pp. 526-532.
- Cardini, F., Wade, C., Regalia, A. L., Gui, S., Li, W., Raschetti, R. & Kronenberg, F. (2006) Clinical research in traditional medicine: priorities and methods, *Complementary therapies in medicine*, 14, pp. 282-287. Doi: 10.1016/j.ctim.2006.07.003.

- Chen, A., Gupte, C., Akhtar, K., Smith, P. & Cobb, J. (2012) The Global Economic Cost of Osteoarthritis: How the UK Compares, *Arthritis*, 2012, pp. 698709-698709. Doi: 10.1155/2012/698709.
- Chen, D. S. & Shun, H. Y. (eds.) (2005). Terminology of Traditional Chinese Medicine, Hunan: Hunan Science and Technology Press.
- Cheng, X. (1999) *Chinese Acupuncture and Moxibustion*. Beijing: Foreign Languages Press.
- Cheon, S., Park, H.-J., Chae, Y. & Lee, H. (2018) Does different information disclosure on placebo control affect blinding and trial outcomes? A case study of participant information leaflets of randomized placebo-controlled trials of acupuncture, *BMC Medical Research Methodology*, 18, pp. 13. Doi: 10.1186/s12874-018-0474-1.
- Cheong, K. B., Zhang, J.-p., Huang, Y. & Zhang, Z.-j. (2013) The effectiveness of acupuncture in prevention and treatment of postoperative nausea and vomiting--a systematic review and meta-analysis, *Plos One*, 8, pp. e82474-e82474. Doi: 10.1371/journal.pone.0082474.
- CICM. (2018) *Acupuncture Degree BSc Hons*: College of Integrated Chinese Medicine. Available from: <https://www.facebook.com/cicmcollege> [Accessed 23 February 2018]
- CNKI. (2017) *User Guide* China Knowledge Resource Integrated Database. Available from: <http://oversea.cnki.net/kns55/UserGuide/en/index.html> [Accessed 22 August 2017]
- Coeytaux, R. R., Chen, W., Lindemuth, C. E., Tan, Y. & Reilly, A. C. (2006) Variability in the diagnosis and point selection for persons with frequent headache by traditional Chinese medicine acupuncturists, *Journal of alternative and complementary medicine*, 12, pp. 863-72. Doi: 10.1089/acm.2006.12.863.
- Colagiuri, B. & Smith, C. A. (2012) A systematic review of the effect of expectancy on treatment responses to acupuncture, *Evidence-Based Complementary And Alternative Medicine: Ecam*, pp. 857804-857804. Doi: 10.1155/2012/857804.
- Colquhoun, D. (2018) *acupuncture – DC's Improbable Science*. Available from: <http://www.dcsience.net/category/acupuncture/> [Accessed 6th June 2018]
- Colquhoun, D. & Novella, S. P. (2013) Acupuncture is theatrical placebo, *Anesthesia And Analgesia*, 116, pp. 1360-1363. Doi: 10.1213/ANE.0b013e31828f2d5e.
- Conaghan, P. (2008) *Osteoarthritis: National clinical guideline for care and management in adults*, Royal College of Physicians (UK).
- Corbett, M. S., Rice, S. J. C., Madurasinghe, V., Slack, R., Fayter, D. A., Harden, M., Sutton, A. J., MacPherson, H. & Woolacott, N. F. (2013) Acupuncture and other physical treatments for the relief of pain due to osteoarthritis of the knee: network meta-analysis, *Osteoarthritis And Cartilage*, 21, pp. 1290-1298. <http://dx.doi.org/10.1016/j.joca.2013.05.007>
- Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I. & Petticrew, M. (2014) *Developing and evaluating complex interventions: new guidance*. Available from: <https://www.mrc.ac.uk/documents/pdf/complex-interventions-guidance/> [Accessed 21 September 2017].
- Creswell, J. W. (2008) *Research design: qualitative, quantitative, and mixed methods approaches*. London: Sage.

- Cui, H. w. 崔红卫 (2013) Wēnzhēnjiǔ zhìliáo xīgǔ guānjiéyán de liáoxiào fēnxī 温针灸治疗膝骨关节炎的疗效分析 [Analysis of the treatment efficacy of warm needle acupuncture therapy for osteoarthritis of the knee], *Zhōngguó yīyào zhǐnán* 中国医药指南 (*Guide of China Medicine*), Number 14, pp 669-670. [Chinese].
- Cummings, M. (2009a) Modellvorhaben Akupunktur--a summary of the ART, ARC and GERAC trials, *Acupuncture In Medicine: Journal Of The British Medical Acupuncture Society*, 27, pp. 26-30. Doi: 10.1136/aim.2008.000281.
- Cummings, M. (2009b) Why recommend acupuncture for low back pain but not for osteoarthritis? A commentary on recent NICE guidelines, *Acupuncture In Medicine: Journal Of The British Medical Acupuncture Society*, 27, pp. 128-129. Doi: 10.1136/aim.2009.001214.
- De Craen, A. J. M., Kaptchuk, T. J., Tijssen, J. G. P. & Kleijnen, J. (1999) Placebos and placebo effects in medicine: Historical overview, *Journal of the Royal Society of Medicine*, 92, pp. 511-515.
- De Vaus, D. A. (2002) *Surveys in social research*. 5th ed. London: Routledge.
- Deadman, P., Al-Khafaji, M. & Baker, K. (2007) *A manual of acupuncture*. Hove, East Sussex: Journal of Chinese Medicine Publications.
- Dhond, R. P., Witzel, T., Hämäläinen, M., Kettner, N. & Napadow, V. (2008a) Spatiotemporal mapping the neural correlates of acupuncture with MEG, *Journal of alternative and complementary medicine*, 14, pp. 679-688. Doi: 10.1089/acm.2007.0824.
- Dhond, R. P., Yeh, C., Park, K., Kettner, N. & Napadow, V. (2008b) Acupuncture modulates resting state connectivity in default and sensorimotor brain networks, *Pain*, 136, pp. 407-418. Doi: 10.1016/j.pain.2008.01.011.
- Ding, J. x. 丁建兴 (2016) Wēn zhēnjiǔ yǔ pǔtōng zhēn cì zhìliáo fēnghán shī xíng xī guānjié gǔ xìng guānjié yán línchuáng xiàoguǒ bǐjiào fēnxī 温针灸与普通针刺治疗风寒湿型膝关节骨性关节炎临床效果比较分析 Comparison of Clinical Effects of Warm Needle Acupuncture and Standard Acupuncture in Treating Wind Cold Damp Type Osteoarthritis of the Knee, *Màn xìng bìng xué zázhi* 慢性病学杂志 (*Chronic Pathematology Journal*), Number 11, pp 1250-1252. [Chinese].
- Dworkin, R. H., Turk, D. C., Wyrwich, K. W., Beaton, D., Cleeland, C. S., Farrar, J. T., Haythornthwaite, J. A., Jensen, M. P., Kerns, R. D., Ader, D. N., Brandenburg, N., Burke, L. B., Cella, D., Chandler, J., Cowan, P., Dimitrova, R., Dionne, R., Hertz, S., Jadad, A. R., Katz, N. P., Kehlet, H., Kramer, L. D., Manning, D. C., McCormick, C., McDermott, M. P., McQuay, H. J., Patel, S., Porter, L., Quessy, S., Rappaport, B. A., Rauschkolb, C., Revicki, D. A., Rothman, M., Schmader, K. E., Stacey, B. R., Stauffer, J. W., von Stein, T., White, R. E., Witter, J. & Zavisic, S. (2008) Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations, *The Journal Of Pain: Official Journal Of The American Pain Society*, 9, pp. 105-121. Doi: 10.1016/j.jpain.2007.09.005.
- Eldridge, S. M., Chan, C. L., Campbell, M. J., Bond, C. M., Hopewell, S., Thabane, L. & Lancaster, G. A. (2016a) CONSORT 2010 statement: extension to randomised pilot and feasibility trials, *Pilot And Feasibility Studies*, 2, pp. 64-64. <https://doi.org/10.1186/s40814-016-0105-8>
- Eldridge, S. M., Lancaster, G. A., Campbell, M. J., Thabane, L., Hopewell, S., Coleman, C. L. & Bond, C. M. (2016b) Defining Feasibility and Pilot Studies in Preparation for

- Randomised Controlled Trials: Development of a Conceptual Framework, *Plos One*, 11, pp. e0150205-e0150205. Doi: 10.1371/journal.pone.0150205.
- Ellis, A., Wiseman, N. & Boss, K. (1991) *Fundamentals of Chinese Acupuncture*. Brookline Massachusetts: Paradigm Publications.
- Elman, B. A. (2011) *Medical Philology in the "Second Rome": Ancient Learning & Attack on "Traditional Chinese Medicine"*. Harvard University. https://www.youtube.com/watch?v=7aJarp7liU0&feature=player_embedded#! [Accessed 5 May 2016]
- Endres, H. G. (2008) Acupuncture: specific and non-specific effects, *Forsch Komplementmed*, 15, pp. 6-8. Doi: 10.1159/000113416.
- Endres, H. G., Molsberger, A. & Haake, M. (2008) Sham acupuncture is not a placebo—reply, *Archives of Internal Medicine*, 168, pp. 1012-1012. Doi: 10.1001/archinte.168.9.1012.
- Ernst, E. & Lee, M. S. (2010) Acupuncture for rheumatic conditions: an overview of systematic reviews, *Rheumatology*, 49, pp. 1957-1961. Doi: rheumatology/keq180.
- Ernst, E. & White, A. R. (1998) Acupuncture for back pain: a meta-analysis of randomized controlled trials, *Archives of Internal Medicine*, 158, pp. 2235-2241. Doi: DOI: 10.1001/archinte.158.20.2235.
- Ezzo, J., Hadhazy, V., Birch, S., Lao, L., Kaplan, G., Hochberg, M. & Berman, B. (2001) Acupuncture for osteoarthritis of the knee: a systematic review, *Arthritis and Rheumatism*, 44, pp. 819-825.
- Fan, X. I. 范小利 (2014) Xún jīngjīn guīlǜ ā shì xué wēnzhenjiǔ zhìliáo xīgǔxìng guānjié yán liáoxiào guānchá 循经筋规律阿是穴温针灸治疗膝骨性关节炎疗效观察 [Observation of the efficacy of the warm needle acupuncture a A-shi points for osteoarthritis of the knee using following Meridian-Muscle- Rules], *Běifāng yàoxué* 北方药学 (*Journal of North Pharmacy*), Number 09, pp 90-91. [Chinese].
- Field, A. P. (2009) *Discovering statistics using SPSS: (and sex and drugs and rock 'n' roll)*. Los Angeles ; London: Sage.
- Filshie, J. & White, A. (1998) *Medical acupuncture: a Western scientific approach*. Edinburgh ; New York: Churchill Livingstone.
- Fischer, F., G. L., Witt, C. M., Linde, K., von Ammon, K., Cardini, F., Falkenberg, T., Fonnebo, V., Johannessen, H., Reiter, B., Uehleke, B., Weidenhammer, W. & Brinkhaus, B. (2014) A research roadmap for complementary and alternative medicine - what we need to know by 2020, *Forschende Komplementärmedizin*, 21, pp. e1–e16. Doi: 10.1159/000360744.
- Fixler, M. (2018) *Courses Japanese Acupuncture London*. Available from: <http://japaneseacupuncturelondon.com/courses> [Accessed 23 February 2018]
- Flick, U. (2009) *An introduction to qualitative research*. Los Angeles ; London: Sage.
- Fønnebo, V., Grimsgaard, S., Walach, H., Ritenbaugh, C., Norheim, A. J., MacPherson, H., Lewith, G., Launsø, L., Koithan, M., Falkenberg, T., Boon, H. & Aickin, M. (2007) Researching complementary and alternative treatments--the gatekeepers are not at home, *BMC Medical Research Methodology*, 7, pp. 7-7. Doi: 10.1186/1471-2288-7-7.
- Foster, N. E., Thomas, E., Barlas, P., Hill, J. C., Young, J., Mason, E. & Hay, E. M. (2007) Acupuncture as an adjunct to exercise based physiotherapy for osteoarthritis of the

- knee: randomised controlled trial, *BMJ Clinical Research*, 335, pp. 436-436. Doi: 10.1136/bmj.39280.509803.BE.
- Fregni, F., Imamura, M., Chien, H. F., Lew, H. L., Boggio, P., Kaptchuk, T. J., Riberto, M., Hsing, W. T., Battistella, L. R. & Furlan, A. (2010) Challenges and recommendations for placebo controls in randomized trials in physical and rehabilitation medicine: a report of the international placebo symposium working group, *American Journal Of Physical Medicine & Rehabilitation / Association Of Academic Physiatrists*, 89, pp. 160-172. Doi: 10.1097/PHM.0b013e3181bc0bbd.
- Frenkel, O. (2008) A phenomenology of the 'placebo effect': taking meaning from the mind to the body, *The Journal of medicine and philosophy*, 33, pp. 58-79. Doi: 10.1093/jmp/jhm005.
- Gale, N. K., Heath, G., Cameron, E., Rashid, S. & Redwood, S. (2013) Using the framework method for the analysis of qualitative data in multi-disciplinary health research, *BMC Medical Research Methodology*, 13, pp. 117. Doi: 10.1186/1471-2288-13-117.
- Ghasemi, A. & Zahediasl, S. (2012) Normality Tests for Statistical Analysis: A Guide for Non-Statisticians, *International Journal of Endocrinology and Metabolism*, 10, pp. 486-489. Doi: 10.5812/ijem.3505.
- Goldacre, B. (2008) *Bad Science*. London: Harper Collins.
- Golding, R. (2013) *The Complete Stems and Branches: Time and Space in Classical Acupuncture*. Moondance Publishers.
- Gotzsche, P. C. (1995) Placebo effects. Concept of placebo should be discarded, *BMJ Clinical Research*, 311, pp. 1640-1641.
- Grant, S. J., Schnyer, R. N., Chang, D. H.-T., Fahey, P. & Bensoussan, A. (2013) Interrater reliability of chinese medicine diagnosis in people with prediabetes, *Evidence-Based Complementary And Alternative Medicine: Ecam*, 2013, pp. 710892-710892. Doi: 10.1155/2013/710892.
- Grünbaum, A. (1986) The placebo concept in medicine and psychiatry, *Psychological Medicine*, 16, pp. 19-38.
- Gu, Y. K. 谷衍奎 (2003) *A Dictionary of the origins of Chinese Characters 汉字源流字典*. Beijing: Hua Xia Publishing House 华夏出版社
- Haake, M., Müller, H.-H., Schade-Brittinger, C., Basler, H. D., Schäfer, H., Maier, C., Endres, H. G., Trampisch, H. J. & Molsberger, A. (2007) German Acupuncture Trials (GERAC) for chronic low back pain: randomized, multicenter, blinded, parallel-group trial with 3 groups, *Archives of Internal Medicine*, 167, pp. 1892-1898.
- Habacher, G., Pittler, M. H. & Ernst, E. (2006) Effectiveness of acupuncture in veterinary medicine: systematic review, *Journal Of Veterinary Internal Medicine*, 20, pp. 480-488.
- Hammerschlag, R. & Stux, G. (eds.) (2001). *Clinical acupuncture: scientific basis*: Berlin; Springer.
- Hansson, S. O. (2014) *Science and Pseudo-Science* Stanford, USA: Stanford: Center for the Study of Language and Information. Available from: <https://plato.stanford.edu/entries/pseudo-science/> [Accessed 11 January 2017]

- Harris, R. E., Zubieta, J. K., Scott, D. J., Napadow, V., Gracely, R. H. & Clauw, D. J. (2009) Traditional Chinese acupuncture and placebo (sham) acupuncture are differentiated by their effects on mu-opioid receptors (MORs), *NeuroImage*, 47, pp. 1077-1085. Doi: 10.1016/j.neuroimage.2009.05.083.
- Hay, E., Barlas, P., Foster, N., Hill, J., Thomas, E. & Young, J. (2004) Is acupuncture a useful adjunct to physiotherapy for older adults with knee pain?: The "Acupuncture, Physiotherapy and Exercise" (APEX) study [ISRCTN88597683], *BMC Musculoskeletal Disorders*, 5, pp. 31. Doi: 10.1186/1471-2474-5-31.
- Hays, R. D., Sherbourne, C. D. & Mazel, R. M. (1993) The rand 36-item health survey 1.0, *Health Economics*, 2, pp. 217-227. Doi: 10.1002/hec.4730020305.
- Health Research Authority. (2014) Health Research Authority: Protecting and promoting the interests of patients and the public in health research. Available from: <http://www.hra.nhs.uk/> [Accessed 10 January 2014]
- Herbert, R. D. & Bø, K. (2005) Analysis of quality of interventions in systematic reviews, *BMJ Clinical Research*, 331, pp. 507-509.
- Hertzog, M. A. (2008) Considerations in determining sample size for pilot studies, *Research in Nursing & Health*, 31, pp. 180-191. Doi: 10.1002/nur.20247.
- Hicks, A., Hicks, J. & Mole, P. (2010) *Five Element Constitutional Acupuncture*. 2 ed. Churchill Livingstone.
- Higgins, J. & Green, S. (2008) *Cochrane handbook for systematic reviews of interventions*: Cochrane Collaboration. Available from: <http://training.cochrane.org/handbook> [Accessed 10 January 2014]
- Hinman, R. S., McCrory, P., Pirotta, M., Relf, I., Crossley, K. M., Reddy, P., Forbes, A., Harris, A., Metcalf, B. R., Kyriakides, M., Novy, K. & Bennell, K. L. (2012) Efficacy of acupuncture for chronic knee pain: protocol for a randomised controlled trial using a Zelen design, *BMC Complementary And Alternative Medicine*, 12, pp. 161-161. Doi: 10.1186/1472-6882-12-161.
- Hinman, R. S., McCrory, P., Pirotta, M., Relf, I., Forbes, A., Crossley, K. M., Williamson, E., Kyriakides, M., Novy, K., Metcalf, B. R., Harris, A., Reddy, P., Conaghan, P. G. & Bennell, K. L. (2014) Acupuncture for chronic knee pain: a randomized clinical trial, *JAMA*, 312, pp. 1313-1322. Doi: 10.1001/jama.2014.12660.
- Hinrichs, T. J. & Barnes, L. L. (2013) *Chinese medicine and healing: an illustrated history*. Cambridge, Mass.: Belknap Press of Harvard University Press.
- Hippisley-Cox, J., Coupland, C. & Logan, R. (2005) Risk of adverse gastrointestinal outcomes in patients taking cyclo-oxygenase-2 inhibitors or conventional non-steroidal anti-inflammatory drugs: population based nested case-control analysis, *BMJ Clinical Research*, 331, pp. 1310-1316.
- Hogeboom, C. J., Sherman, K. J. & Cherkin, D. C. (2001) Variation in diagnosis and treatment of chronic low back pain by traditional Chinese medicine acupuncturists, *Complementary therapies in medicine*, 9, pp. 154-166. Doi: 10.1054/ctim.2001.0457,.
- Hopton, A. & MacPherson, H. (2010) Acupuncture for chronic pain: is acupuncture more than an effective placebo? A systematic review of pooled data from meta-analyses, *Pain Practice: The Official Journal Of World Institute Of Pain*, 10, pp. 94-102. Doi: 10.1111/j.1533-2500.2009.00337.x.

- Hopton, A. K., Curnoe, S., Kanaan, M. & MacPherson, H. (2012) Acupuncture in practice: mapping the providers, the patients and the settings in a national cross-sectional survey, *BMJ Open*, 2, pp. 1-9. Doi: 10.1136/bmjopen-2011-000456.
- Hou, P.-W., Fu, P.-K., Hsu, H.-C. & Hsieh, C.-L. (2015) Traditional Chinese medicine in patients with osteoarthritis of the knee, *Journal Of Traditional And Complementary Medicine*, 5, pp. 182-196. Doi: 10.1016/j.jtcme.2015.06.002.
- Hróbjartsson, A. (2002) What are the main methodological problems in the estimation of placebo effects?, *Journal of clinical epidemiology*, 55, pp. 430-435.
- Hróbjartsson, A. (2008) Clinical placebo interventions are unethical, unnecessary, and unprofessional, *The Journal Of Clinical Ethics*, 19, pp. 66-69.
- Hróbjartsson, A. & Gøtzsche, P. C. (2010) Placebo interventions for all clinical conditions, *Cochrane Database of Systematic Reviews*. Issue 3. Available from: <http://dx.doi.org/10.1002/14651858.CD003974.pub3> [Accessed September 21 2017].
- Hsu, C., Sherman, K. J., Eaves, E. R., Turner, J. A., Cherkin, D. C., Crompton, D., Schafer, L. & Ritenbaugh, C. (2014) New perspectives on patient expectations of treatment outcomes: results from qualitative interviews with patients seeking complementary and alternative medicine treatments for chronic low back pain, *BMC Complementary And Alternative Medicine*, 14, pp. 276-276. Doi: 10.1186/1472-6882-14-276.
- Hu, J. h. 胡江红 (2015) Wēn zhēnjiǔ zài hán shī xíng xī guānjié gǔ xìng guānjié yán zhōng de liáoxiào fēnxī 温针灸在寒湿型膝关节骨性关节炎中的疗效分析 [Analysis of curative effect of warm needle acupuncture on Cold Damp type osteoarthritis of the knee], *Zhēnjiǔ líncuáng zázhi* 针灸临床杂志 (*Journal of Clinical Acupuncture and Moxibustion*), Number 08, pp 30-31. [Chinese].
- Hu, N. n., Liu, C. h. & Chun, H. j. 胡妮娜, 刘春红 & 陈会君 (2014) Wēnzhēnjiǔ zhìliáo xīguānjié gǔxìng guānjiéyán liáoxiào guānchá 温针灸治疗膝关节骨性关节炎疗效观察 [Observation of the treatment efficacy of warm needle acupuncture therapy for osteoarthritis of the knee], *Yàowù yǔ rén* 药物与人 (*Medicine and People*), Number 05, pp 174-175. [Chinese].
- Hua, B., Abbas, E., Hayes, A., Ryan, P., Nelson, L. & O'Brien, K. (2012) Reliability of Chinese medicine diagnostic variables in the examination of patients with osteoarthritis of the knee, *Journal of alternative and complementary medicine*, 18, pp. 1028-1037. Doi: 10.1089/acm.2011.0621.
- Huang, C. x. 黄朝曦 (2007) Wēnzhēnjiǔ zhìliáo xīguānjié gǔxìng guānjiéyán 52 lì 温针灸治疗膝关节骨性关节炎52例 [Warm needle acupuncture therapy for 52 cases of osteoarthritis of the knee], *Fújiàn zhōngyīyào* 福建中医药 (*Fujian Journal of Traditional Chinese Medicine*), Number 05, pp 23. [Chinese].
- Huang, W., Pach, D., Napadow, V., Park, K., Long, X., Neumann, J., Maeda, Y., Nierhaus, T., Liang, F. & Witt, C. M. (2012) Characterizing acupuncture stimuli using brain imaging with fMRI--a systematic review and meta-analysis of the literature, *Plos One*, 7, pp. e32960. Doi: 10.1371/journal.pone.0032960.
- Huangfu, M. 皇甫谧 (282) 针灸甲乙经 *The AB Classic of Acupuncture and Moxibustion*, 中医经典诵读丛书 The Chinese medicine recitation classics series, Published (2005) 人民军医出版社 People's Military Medical Press, 北京 Beijing. [Chinese].

- Jakes, D., Kirk, R. & Muir, L. (2014) A qualitative systematic review of patients' experiences of acupuncture, *Journal of alternative and complementary medicine*, 20, pp. 663-671. Doi: 10.1089/acm.2013.0446.
- Jarrett, L. (2006) *The Clinical Practice of Chinese Medicine*. Spirit Path Press.
- Jarrett, L. S. (1999) *Nourishing Destiny: The Inner Tradition of Chinese Medicine*. Spirit Path Press.
- Kalauokalani, D., Cherkin, D. C., Sherman, K. J., Koepsell, T. D. & Deyo, R. A. (2001a) Lessons from a trial of acupuncture and massage for low back pain: patient expectations and treatment effects, *Spine*, 26, pp. 1418-1424.
- Kalauokalani, D., Sherman, K. J. & Cherkin, D. C. (2001b) Acupuncture for chronic low back pain: diagnosis and treatment patterns among acupuncturists evaluating the same patient, *Southern Medical Journal*, 94, pp. 486-492.
- Kaptchuk, T. J. (2003) Effect of interpretive bias on research evidence, *BMJ Clinical Research*, 326, pp. 1453-1455.
- Kaptchuk, T. J., Chen, K. & Song, J. (2010) Recent clinical trials of acupuncture in the West: responses from the practitioners, *Chinese Journal of Integrative Medicine*, 16, pp. 197-203. <http://dx.doi.org/10.1007/s11655-010-0197-x>
- Kaptchuk, T. J., Kelley, J. M., Conboy, L. A., Davis, R. B., Kerr, C. E., Jacobson, E. E., Kirsch, I., Schyner, R. N., Nam, B. H., Nguyen, L. T., Park, M., Rivers, A. L., McManus, C., Kokkotou, E., Drossman, D. A., Goldman, P. & Lembo, A. J. (2008) Components of placebo effect: randomised controlled trial in patients with irritable bowel syndrome, *BMJ Clinical Research*, 336, pp. 999-1003. Doi: 10.1136/bmj.39524.439618.25.
- Kaptchuk, T. J., Shaw, J., Kerr, C. E., Conboy, L. A., Kelley, J. M., Csordas, T. J., Lembo, A. J. & Jacobson, E. E. (2009) "Maybe I made up the whole thing": placebos and patients' experiences in a randomized controlled trial, *Culture, medicine and psychiatry*, 33, pp. 382-411. Doi: 10.1007/s11013-009-9141-7.
- Kaptchuk, T. J., Stason, W. B., Davis, R. B., Legedza, A. R. T., Schnyer, R. N., Kerr, C. E., Stone, D. A., Nam, B. H., Kirsch, I. & Goldman, R. H. (2006) Sham device v inert pill: randomised controlled trial of two placebo treatments, *BMJ Clinical Research*, 332, pp. 391-397. Doi: 10.1136/bmj.38726.603310.55.
- Karner, M., Brazkiewicz, F., Remppis, A., Fischer, J., Gerlach, O., Stremmel, W., Subramanian, S. V. & Greten, H. J. (2013) Objectifying specific and nonspecific effects of acupuncture: a double-blinded randomised trial in osteoarthritis of the knee, *Evidence-Based Complementary And Alternative Medicine: Ecam*, 2013, pp. 427265-427265. Doi: 10.1155/2013/427265.
- Kavoussi, B. (2015) Acupuncture for low back pain: ritual healing or medicine?, *Focus on Alternative and Complementary Therapies*, 20, pp. 2-12. Doi: 10.1111/fct.12161.
- Kim, M., Cobbin, D. & Zaslowski, C. (2008a) Traditional Chinese medicine tongue inspection: an examination of the inter- and intrapractitioner reliability for specific tongue characteristics, *Journal of alternative and complementary medicine*, 14, pp. 527-536. Doi: 10.1089/acm.2007.0079.
- Kim, Y., Park, J., Lee, H., Bang, H. & Park, H.-J. (2008b) Content validity of an acupuncture sensation questionnaire, *Journal of alternative and complementary medicine*, 14, pp. 957-963. Doi: 10.1089/acm.2007.0565.

- King, E., Cobbin, D., Walsh, S. & Ryan, D. (2002) The reliable measurement of radial pulse characteristics, *Acupuncture In Medicine: Journal Of The British Medical Acupuncture Society*, 20, pp. 150-159.
- King, E., Walsh, S. & Cobbin, D. (2006) The testing of classical pulse concepts in Chinese medicine: left- and right-hand pulse strength discrepancy between males and females and its clinical implications, *Journal of alternative and complementary medicine*, 12, pp. 445-450.
- Kong, J., Kaptchuk, T. J., Polich, G., Kirsch, I., Vangel, M., Zyloney, C., Rosen, B. & Gollub, R. (2009a) Expectancy and treatment interactions: a dissociation between acupuncture analgesia and expectancy evoked placebo analgesia, *NeuroImage*, 45, pp. 940-949. Doi: 10.1016/j.neuroimage.2008.12.025.
- Kong, J., Kaptchuk, T. J., Polich, G., Kirsch, I., Vangel, M., Zyloney, C., Rosen, B. & Gollub, R. L. (2009b) An fMRI study on the interaction and dissociation between expectation of pain relief and acupuncture treatment, *NeuroImage*, 47, pp. 1066-1076. Doi: 10.1016/j.neuroimage.2009.05.087.
- Kumar, P. & Clark, M. L. (2017) Kumar & Clark's clinical medicine. 9th ed. Edinburgh: Elsevier. [i-book]
- Langevin, H., Hammerschlag, R., Lao, L., Napadow, V., Schnyer, R. N. & Sherman, K. J. (2006) Controversies in acupuncture research: selection of controls and outcome measures in acupuncture clinical trials, *Journal of alternative and complementary medicine*, 12, pp. 943-953.
- Langevin, H. M., Wayne, P. M., Hugh, M., Schnyer, R., Milley, R. M., Napadow, V., Lixing, L., Park, J., Harris, R. E., Cohen, M., Sherman, K. J., Haramati, A. & Hammerschlag, R. (2011) Paradoxes in Acupuncture Research: Strategies for Moving Forward, *Evidence-based Complementary & Alternative Medicine (eCAM)*, 8, pp. 1-11. Doi: 2011/180805.
- Lansdown, H., Howard, K., Brealey, S. & MacPherson, H. (2009) Acupuncture for pain and osteoarthritis of the knee: a pilot study for an open parallel-arm randomised controlled trial, *BMC Musculoskeletal Disorders*, 10, pp. 130-130.
- Lao, L., Ezzo, J., Berman Brian, M. & Hammerschlag, R. (2001). Assessing clinical efficacy of acupuncture: Considerations for designing future acupuncture trials. In: Hammerschlag, R. & Stux, G. (eds.) *Clinical acupuncture: scientific basis*. Berlin: Springer, pp. 187-209.
- Laucis, N. C., Hays, R. D. & Bhattacharyya, T. (2015) Scoring the SF-36 in Orthopaedics: A Brief Guide, *The Journal of Bone And Joint Surgery*, 97, pp. 1628-1634. Doi: 10.2106/jbjs.o.00030.
- Law Simon, K. & Li, T. (2013) Acupuncture for glaucoma, *Cochrane Database of Systematic Reviews*. Issue 5. Available from: <http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD006030.pub3/full> [Accessed 23 October 2017].
- Lee, A., Chan, S. K. C. & Fan, L. T. Y. (2015) Stimulation of the wrist acupuncture point PC6 for preventing postoperative nausea and vomiting, *Cochrane Database of Systematic Reviews*. Issue 11. Available from: <http://dx.doi.org/10.1002/14651858.CD003281.pub4> [Accessed 20 September 2017].
- Lee, J., Napadow, V. & Park, K. (2014) Pain and sensory detection threshold response to acupuncture is modulated by coping strategy and acupuncture sensation, *BMC Complementary And Alternative Medicine*, 14, pp. 324-324.

- Lee, M. S. & Ernst, E. (2011) Acupuncture for pain: an overview of Cochrane reviews, *Chinese Journal of Integrative Medicine*, 17, pp. 187-189. Doi: 10.1007/s11655-011-0665-7.
- Legge, D. & Vance, K. (1997) Close to the bone: the treatment of musculo-skeletal disorder with acupuncture and other traditional Chinese medicine. Woy Woy, N.S.W.: Sydney College Press.
- Lewith, G. & Vincent, C. (1995) Evaluation of the clinical effects of acupuncture: A problem reassessed and a framework for future research, *Pain Forum*, 4, pp. 29-39. Doi: [https://doi.org/10.1016/S1082-3174\(11\)80073-5](https://doi.org/10.1016/S1082-3174(11)80073-5).
- Li, C. d., Huang, X. y., Yang, X. g., Wang, Q. f. & Huang, S. q. 李常度, 黄信勇, 杨旭光, 王琼芬 & 黄思琴 (2006) Wēnzhēnjiǔ zhìliáo xūhánxíng xīgǔ guānjiéyán liáoxiào guānchá 温针灸治疗虚寒型膝关节炎疗效观察 [Observation of the treatment efficacy of all needle acupuncture therapy for deficient cold type osteoarthritis of the knee], *Zhōngguó zhēnjiǔ 中国针灸 (Chinese Acupuncture and Moxibustion)*, Number 03, pp 189-191. [Chinese].
- Li, L., Hong, K. d., Wu, M. x. & He, F. 李俐, 洪昆达, 吴明霞 & 何锋 (2011) Wēnzhēnjiǔ zhìliáo yángxūhán níng xíng xīgǔxíng guānjiéyán 40 lì 温针灸治疗阳虚寒凝型膝骨性关节炎40例 [Warm needle acupuncture therapy for 40 cases of Yang deficiency congealing cold type osteoarthritis of the knee], *Zhōngguó zhēnjiǔ xuéhuì niánhuì 中国针灸学会年会 (China Association of Acupuncture and Moxibustion Annual Conference)*, pp 4. [Chinese].
- Li, Z. h. & Li, D. m. 李志宏 & 李冬梅 (2010) Wēnzhēnjiǔ zhìliáo xīgūguānjié gǔxíng guānjiéyán 55 lì 温针灸治疗膝关节骨性关节炎55例 [Warm needle acupuncture therapy for 55 cases of osteoarthritis of the knee], *Zhōngyī wàizhì zázhi 中医外治杂志 (Journal of External Therapy of Traditional Chinese Medicine)*, Number 05, pp 32-33. [Chinese].
- Lin, X., Huang, K., Zhu, G., Huang, Z., Qin, A. & Fan, S. (2016) The Effects of Acupuncture on Chronic Knee Pain Due to Osteoarthritis: A Meta-Analysis, *The Journal Of Bone And Joint Surgery. American Volume*, 98, pp. 1578-1585. Doi: 10.2106/JBJS.15.00620.
- Linde, K., Witt, C. M., Streng, A., Weidenhammer, W., Wagenpfeil, S., Brinkhaus, B., Willich, S. N. & Melchart, D. (2007) The impact of patient expectations on outcomes in four randomized controlled trials of acupuncture in patients with chronic pain, *Pain*, 128, pp. 264-271. Doi: 10.1016/j.pain.2006.12.006.
- Liu, B. 刘博 (2016) Wēn zhēnjiǔ zhìliáo xī guānjié gǔ guānjié yán de liáoxiào fēnxī 温针灸治疗膝关节骨性关节炎的疗效分析 Analysis of Curative Effect of Warm Needle Acupuncture on Knee Joint Osteoarthritis, *Zhōngguó wèishēng biāozhǔn guǎnlǐ 中国卫生标准管理 [China Health Standards Management]*, Number 03, pp 135-136. [Chinese].
- Liu, F. & Du, J. 刘峰 & 杜炯 (2015) Wēn zhēnjiǔ zhìliáo xī gǔ xíng guānjié yán de línchuáng yánjiū 温针灸治疗膝骨性关节炎的临床研究 [Clinical Study on Treatment of Knee Osteoarthritis with Warm Needle Technique], *Zhōng yīyào línchuáng zázhi 中医药临床杂志 (Clinical Journal of Traditional Chinese Medicine)*, Number 08, pp 1076-1079. [Chinese].

- Liu, L. a., Ma, C. y. & Jiang, W. 刘立安, 马春燕 & 姜文 (2003) Wēnzhēnjiǔ zhìliáo lǎoniánxìng xīgǔ guānjiébing de línchuáng guānchá 温针灸治疗老年性膝骨关节病的临床观察 (A clinical observation on treatment of senile osteoarthopathy by warm needle moxibustion), *Zhōngguó zhēnjiǔ* 中国针灸 (Chinese Acupuncture and Moxibustion), Number 10, pp 15-16. [Chinese].
- Liu, T. (2009) Acupuncture: what underlies needle administration?, *Evidenced Based Complementary and Alternative Medicine*, 6, pp. 185-93. Doi: 10.1093/ecam/nen002.
- Lloyd, G. & Sivin, N. (2002) *The Way and the Word: Science and Medicine in early China and Greece*. New Haven and London: Yale University Press.
- LSBU (2011) *Code of Practice for Research Involving Human Participants*, London South Bank University Ethics Committee London: <https://www1.lsbu.ac.uk/staff/academic/research/docs/ethics-code-of-practice.pdf>.
- Lu, X. h. 卢兴华 (2015) Wēn zhēnjiǔ zhìliáo lǎonián xī guānjié gǔ xìng guānjié yán 81 lì línchuáng guānchá 温针灸治疗老年膝关节骨性关节炎81例临床观察 [Clinical Observation on 81 Cases of Old-aged Knee Joint Osteoarthritis Treated by Warm Needle Acupuncture], *Zhōngguó mínzú mínjiān yīyào* 中国民族民间医药 (Chinese Journal of Ethnomedicine and Ethnopharmacy), Number 08, pp 62. [Chinese].
- Luo, J., Xu, H. & Liu, B. (2015) Real world research: a complementary method to establish the effectiveness of acupuncture, *BMC Complementary And Alternative Medicine*, 15, pp. 153. Doi: 10.1186/s12906-015-0676-6.
- Luo, X. f. & Mao, H. r. 罗新福 & 毛红蓉 (2011) Wēnzhēn zhìliáo tuìxíngxìng xīguānjiébing línchuáng guānchá 温针治疗退行性膝关节病临床观察 [A clinical observation of the warm needle therapy for degenerative knee joint disease], *Húběi zhōngyī zázhi* 湖北中医杂志 (Hubei Journal of Traditional Chinese Medicine), Number 07, pp 65-66. [Chinese].
- Lv, M. 吕梅 (2000) Wēnzhēn zhìliáo xīguānjié gǔxìng guānjiéyán 48 lì 温针治疗膝关节骨性关节炎48例 [Warm needle therapy for 48 cases of osteoarthritis of the knee], *Zhōngguó zhēnjiǔ* 中国针灸 (Chinese Acupuncture and Moxibustion), Number S1, pp 28-29. [Chinese].
- Lysholm, J. & Gillquist, J. (1982) Evaluation of knee ligament surgery results with special emphasis on use of a scoring scale, *The American Journal Of Sports Medicine*, 10, pp. 150-154.
- Ma, H. b. 马洪滨 (2016) Wēn zhēnjiǔ zhìliáo xī guānjié gǔ guānjié yán de xiàoguǒ yánjiū 温针灸治疗膝关节骨关节炎的效果研究 [Effect of Warm Needle Acupuncture on Knee Joint Osteoarthritis], *Zhōngguó wèishēng biāozhǔn guǎnlǐ* 中国卫生标准管理 [China Health Standards Management], Number 17, pp 135-136. [Chinese].
- Ma, Y.-t., Cho, Z. H. & Ma, M. (2005) *Biomedical acupuncture for pain management: an integrative approach*. St. Louis, Mo.: Elsevier Churchill Livingstone.
- Maciocia, G. (1994) *The practice of Chinese medicine: the treatment of diseases with acupuncture and Chinese herbs*. Edinburgh: Churchill Livingstone.
- Maciocia, G. (2005) *The foundations of Chinese medicine: a comprehensive text for acupuncturists and herbalists*. Edinburgh: Elsevier Churchill Livingstone.

- Maciocia, G. (2006) *The Channels of Acupuncture: Clinical Use of the Secondary Channels and Eight Extraordinary Vessels*. 1st ed. Churchill Livingstone.
- MacPherson, H., Altman, D. G., Hammerschlag, R., Youping, L., Taixiang, W., White, A. & Moher, D. (2010) Revised STAndards for Reporting Interventions in Clinical Trials of Acupuncture (STRICTA): extending the CONSORT statement, *Journal of alternative and complementary medicine*, 16, pp. ST1-ST14. Doi: 10.1089/acm.2010.1610.
- MacPherson, H. & Hammerschlag, R. (2012) Acupuncture and the emerging evidence base: contrived controversy and rational debate, *Journal Of Acupuncture And Meridian Studies*, 5, pp. 141-147. Doi: 10.1016/j.jams.2012.05.001.
- MacPherson, H., Hammerschlag, R., Coeytaux, R. R., Davis, R. T., Harris, R. E., Kong, J.-T., Langevin, H. M., Lao, L., Milley, R. J., Napadow, V., Schnyer, R. N., Stener-Victorin, E., Witt, C. M. & Wayne, P. M. (2016) Unanticipated Insights into Biomedicine from the Study of Acupuncture, *Journal of alternative and complementary medicine*, 22, pp. 101-107. Doi: 10.1089/acm.2015.0184.
- MacPherson, H., Hammerschlag, R., Lewith, G. & Schnyer, R. (2007) *Acupuncture research: strategies for establishing an evidence base*. Edinburgh ; New York: Churchill Livingstone Elsevier.
- MacPherson, H., Nahin, R., Paterson, C., Cassidy, C. M., Lewith, G. T. & Hammerschlag, R. (2008) Developments in acupuncture research: big-picture perspectives from the leading edge, *Journal of alternative and complementary medicine*, 14, pp. 883-887.
- MacPherson, H. & Thomas, K. (2005) Short term reactions to acupuncture--a cross-sectional survey of patient reports, *Acupuncture In Medicine: Journal Of The British Medical Acupuncture Society*, 23, pp. 112-120.
- MacPherson, H., Thomas, K., Walters, S. & Fitter, M. (2001) The York acupuncture safety study: prospective survey of 34 000 treatments by traditional acupuncturists, *BMJ Clinical Research*, 323, pp. 486-487.
- MacPherson, H., Thorpe, L. & Thomas, K. (2006) Beyond needling--therapeutic processes in acupuncture care: a qualitative study nested within a low-back pain trial, *Journal of alternative and complementary medicine*, 12, pp. 873-880.
- MacPherson, H., Thorpe, L., Thomas, K. & Campbell, M. (2004) Acupuncture for low back pain: traditional diagnosis and treatment of 148 patients in a clinical trial, *Complementary therapies in medicine*, 12, pp. 38-44.
- MacPherson, H., Vertosick, E., Lewith, G., Linde, K., Sherman, K. J., Witt, C. M. & Vickers, A. J. (2014) Influence of control group on effect size in trials of acupuncture for chronic pain: a secondary analysis of an individual patient data meta-analysis, *Plos One*, 9, pp. e93739-e93739. Doi: 10.1371/journal.pone.0093739.
- Madsen, M. V., Gøtzsche, P. C. & Hróbjartsson, A. (2009) Acupuncture treatment for pain: systematic review of randomised clinical trials with acupuncture, placebo acupuncture, and no acupuncture groups, *BMJ Clinical Research*, 338, pp. a3115-a3115. Doi: 10.1136/bmj.a3115.
- Manheimer, E., Cheng, K., Linde, K., Lao, L., Yoo, J., Wieland, S., van der Windt, D. A. W. M., Berman, B. M. & Bouter, L. M. (2010) Acupuncture for peripheral joint osteoarthritis, *Cochrane Database of Systematic Reviews*. Issue 1. Available from: <http://dx.doi.org/10.1002/14651858.CD001977.pub2> [Accessed 23 October 2017].

- Manheimer, E., Linde, K., Lao, L., Bouter, L. M. & Berman, B. M. (2007) Meta-analysis: acupuncture for osteoarthritis of the knee, *Annals of Internal Medicine*, 146, pp. 868-877.
- Manyanga, T., Froese, M., Zarychanski, R., Abou-Setta, A., Friesen, C., Tennenhouse, M. & Shay, B. L. (2014) Pain management with acupuncture in osteoarthritis: a systematic review and meta-analysis, *BMC Complementary And Alternative Medicine*, 14, pp. 312-312. Doi: 10.1186/1472-6882-14-312.
- Mason, J. (2002) Qualitative researching. London: Sage.
- Matsumoto, K. & Euler, D. (2008) Kiiko Matsumoto's Clinical Strategies: In the Spirit of Master Nagano. Natwick, MA 01760: Kiiko Matsumoto International.
- Mavrommatis, C. I., Argyra, E., Vadalouka, A. & Vasilakos, D. G. (2012) Acupuncture as an adjunctive therapy to pharmacological treatment in patients with chronic pain due to osteoarthritis of the knee: a 3-armed, randomized, placebo-controlled trial, *Pain*, 153, pp. 1720-1726. Doi: 10.1016/j.pain.2012.05.005.
- Moerman, D. (2002) Meaning, Medicine and the 'Placebo Effect'. Cambridge University Press.
- Moerman, D. E. (2013) Against the "placebo effect": a personal point of view, *Complementary therapies in medicine*, 21, pp. 125-130. Doi: DOI: 10.1016/j.ctim.2013.01.005.
- Moerman, D. E. & Jonas, W. B. (2002) Deconstructing the placebo effect and finding the meaning response, *Annals of Internal Medicine*, 136, pp. 471-476.
- Moffet, H. H. (2009) Sham acupuncture may be as efficacious as true acupuncture: a systematic review of clinical trials, *Journal of alternative and complementary medicine*, 15, pp. 213-216. Doi: 10.1089/acm.2008.0356.
- Moher, D., Hopewell, S., Schulz, K. F., Montori, V., Gøtzsche, P. C., Devereaux, P. J., Elbourne, D., Egger, M. & Altman, D. G. (2010) CONSORT 2010 Explanation and Elaboration: Updated guidelines for reporting parallel group randomised trials, *Journal of clinical epidemiology*, 63, pp. e1-e37.
- Molsberger, A. F., Streitberger, K., Kraemer, J., Brittinger, C. S., Witte, S., Boewing, G. & Haake, M. (2006) Designing an acupuncture study: II. The nationwide, randomized, controlled German acupuncture trials on low-back pain and gonarthrosis, *Journal of alternative and complementary medicine*, 12, pp. 733-742.
- Murray, C. J. L., Vos, T., Lozano, R., Naghavi, M. & et al. (2012) Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010, *The Lancet*, 380, pp. 2197-2223. [http://dx.doi.org/10.1016/S0140-6736\(12\)61689-4](http://dx.doi.org/10.1016/S0140-6736(12)61689-4)
- Napadow, V., Ahn, A., Longhurst, J., Lao, L., Stener-Victorin, E., Harris, R. & Langevin, H. M. (2008) The status and future of acupuncture mechanism research, *Journal of alternative and complementary medicine*, 14, pp. 861-869. Doi: 10.1089/acm.2008.SAR-3.
- National Institute for Health Research. (2014) *NIHR Evaluation, Trials and Studies | Home*. Available from: <https://www.nihr.ac.uk/> [Accessed 10 January 2014]
- Neogi, T. (2013) The epidemiology and impact of pain in osteoarthritis, *Pain in Osteoarthritis*, 21, pp. 1145-1153. <http://dx.doi.org/10.1016/j.joca.2013.03.018>
- Neuman, W. L. (2009) Social Research Methods: Qualitative and Quantitative Approaches. 7 ed. Pearson.

- NICE. (2014a) *BNF: British National Formulary: The National Institute for Health and Care Excellence*. Available from: <https://bnf.nice.org.uk/drug/nimodipine.html> [Accessed 26 April 2018]
- NICE. (2014b) *Osteoarthritis: care and management: National Institute for Health and Clinical Excellence*. Available from: <https://www.nice.org.uk/guidance/cg177> [Accessed Report CG177]
- Novella, S. (2010) *Acupuncture Pseudoscience in the New England Journal of Medicine*. Available from: <https://www.sciencebasedmedicine.org/tag/nejm/> [Accessed 20 December 2016]
- O'Brien, K. A., Abbas, E., Movsessian, P., Hook, M., Komesaroff, P. A. & Birch, S. (2009a) Investigating the reliability of Japanese toyohari meridian therapy diagnosis, *Journal of alternative and complementary medicine*, 15, pp. 1099-1105. Doi: 10.1089/acm.2009.0020.
- O'Brien, K. A., Abbas, E., Zhang, J., Guo, Z.-X., Luo, R., Bensoussan, A. & Komesaroff, P. A. (2009b) An investigation into the reliability of Chinese medicine diagnosis according to Eight Guiding Principles and Zang-Fu Theory in Australians with hypercholesterolemia, *Journal of alternative and complementary medicine*, 15, pp. 259-266. Doi: 10.1089/acm.2008.0204.
- O'Brien, K. A., Abbas, E., Zhang, J., Guo, Z. X., Luo, R., Bensoussan, A. & Komesaroff, P. A. (2009c) Understanding the reliability of diagnostic variables in a Chinese Medicine examination, *Journal of alternative and complementary medicine*, 15, pp. 727-34. Doi: 10.1089/acm.2008.0554.
- O'Brien, K. A. & Birch, S. (2009) A review of the reliability of traditional East Asian medicine diagnoses, *Journal of complementary and alternative medicine*, 15(4), pp. 353-66. Doi: 10.1089/acm.2008.0455.
- O'Connell, N. E., Wand, B. M. & Goldacre, B. (2009) Interpretive bias in acupuncture research?: A case study, *Evaluation & the health professions*, 32, pp. 393-409. Doi: 10.1177/0163278709353394.
- O'Connor, J. & Bensky, D. (eds.) (1981). *Acupuncture : a comprehensive text* - Shanghai College of Traditional Medicine, Chicago: Eastland Press.
- Ortiz, R., Chandros Hull, S. & Colloca, L. (2016) Patient attitudes about the clinical use of placebo: qualitative perspectives from a telephone survey, *BMJ Open*, 6, pp. e011012-e011012. Doi: 10.1136/bmjopen-2015-011012.
- Oxford Dictionary of English (2005) 2nd Edition. Oxford and New York: Oxford University Press
- Oxford Living Dictionaries. (2018) *tonify* | Definition of tonify in English by Oxford Dictionaries: Oxford University Press. Available from: <https://en.oxforddictionaries.com/definition/tonify> [Accessed 21 February 2018]
- Parahoo, K. (1997) *Nursing Research: Principles, Process and Issues*. Great Britain: Macmillan press Ltd.
- Paterson, C., Baarts, C., Launsø, L. & Verhoef, M. J. (2009) Evaluating complex health interventions: a critical analysis of the 'outcomes' concept, *BMC Complementary And Alternative Medicine*, 9, pp. 18-18. Doi: 10.1186/1472-6882-9-18.

- Paterson, C. & Britten, N. (2004) Acupuncture as a complex intervention: a holistic model, *Journal of alternative and complementary medicine*, 10, pp. 791-801. Doi: 10.1089/acm.2004.10.791.
- Paterson, C., Zheng, Z., Xue, C. & Wang, Y. (2008) "Playing their parts": the experiences of participants in a randomized sham-controlled acupuncture trial, *Journal of alternative and complementary medicine*, 14, pp. 199-208. Doi: 10.1089/acm.2007.0682.
- Peng, W., Wang, Y., Zhang, Y. & Liang, C. M. (2007) Acupuncture for vascular dementia, *Cochrane Database of Systematic Reviews*. Issue 2. Available from: <http://dx.doi.org/10.1002/14651858.CD004987.pub2> [Accessed 23 October 2017].
- Petrovic, P., Kalso, E., Petersson, K. M. & Ingvar, M. (2002) Placebo and opioid analgesia-- imaging a shared neuronal network, *Science*, 295, pp. 1737-1740.
- Pickering, A. (2010) *The Mangle of Practice Time, Agency and Science*. University of Chicago Press. [iBook]
- Prady, S. L., Burch, J., Vanderbloemen, L., Crouch, S. & MacPherson, H. (2015) Measuring expectations of benefit from treatment in acupuncture trials: a systematic review, *Complementary therapies in medicine*, 23, pp. 185-199. Doi: 10.1016/j.ctim.2015.01.007.
- Purepong, N., Jitvimonrat, A., Sitthipornvorakul, E., Eksakulkla, S. & Janwantanakul, P. (2012) External validity in randomised controlled trials of acupuncture for osteoarthritis knee pain, *Acupuncture in Medicine*, 30, pp. 187-194. Doi: 10.1136/acupmed-2012-010140.
- Qin, D. 秦达 (2016) Wēn zhēnjiǔ yǔ zhēn cì zhīliáo lǎonián xìng xī gǔ guānjié bìng de línchuáng liáoxiào guānchá 温针灸与针刺治疗老年性膝骨关节病的临床疗效观察 Clinical Observation on Treatment of Old Aged Knee Osteoarthritis by Warm Needle Technique and Needle Acupuncture, *Zhōngguó zhōngyào zázhi* 中国中药杂志 (*China Journal of Chinese Materia Medica*), pp 2. [Chinese].
- Quah, S. R. (2003) Traditional healing systems and the ethos of science, *Social science & medicine*, 57, pp. 1997-2012. Doi: 10.1016/S0277-9536(03)00078-9.
- RAND. (No date-a) *Medical Outcomes Study: Measures of Quality of Life Core Survey from RAND Health*. Available from: http://www.rand.org/health/surveys_tools/mos.html [Accessed 3 April 2014]
- RAND. (no date-b) *Scoring Rules for the RAND 36-Item Health Survey (Version 1.0): Rand Health*. Available from: https://www.rand.org/health/surveys_tools/mos/36-item-short-form.html [Accessed 30 August 2016]
- Ritchie, J. & Spencer, L. (1995). Qualitative data analysis for applied policy research. In: Bryman, A. & Burgess, R. G. (eds.) *Analyzing qualitative data*. London: Routledge, pp. 173-194.
- Rosenberg, A. (2005) *Philosophy of science: a contemporary introduction*. New York: Routledge.
- Salzberg, S. (2009) *Stimulus funds are promoting pseudoscience at Harvard Medical School*. Available from: <http://genome.fieldofscience.com/2009/11/stimulus-funds-are-promoting.html> [Accessed 20 November 2016]
- Salzberg, S. (2016) *Washington Post's Science Section Descends Into Pseudoscience About Acupuncture* USA: Forbes Media LLC. Available from: <http://www.forbes.com/>

- sites/stevensalzberg/2016/05/09/washington-posts-science-section-descends-into-pseudoscience-about-acupuncture/#32368a457c07 [Accessed 20 December 2016]
- Sapolsky, R. M. (2004) Why zebras don't get ulcers: the acclaimed guide to stress, stress-related diseases, and coping. New York: St. Martin's Griffin.
- Scharf, H.-P., Mansmann, U., Streitberger, K., Witte, S., Krämer, J., Maier, C., Trampisch, H.-J. & Victor, N. (2006) Acupuncture and knee osteoarthritis: a three-armed randomized trial, *Annals of Internal Medicine*, 145, pp. 12-20.
- Scheid, V. (2002) Chinese medicine in contemporary China: plurality and synthesis. Durham, NC: Duke University Press.
- Scheid, V. (2007) Traditional Chinese medicine--what are we investigating? The case of menopause, *Complementary therapies in medicine*, 15, pp. 54-68. Doi: 10.1016/j.ctim.2005.12.002.
- Scheid, V. (2008) Authenticity, best practice, and the evidence mosaic. The challenge of integrating traditional East Asian medicines into Western health care, *Complementary therapies in medicine*, 16, pp. 107-108. Doi: 10.1016/j.ctim.2008.03.002.
- Scheid, V. & MacPherson, H. (2012) Integrating East Asian medicine into contemporary healthcare. Edinburgh ; New York: Churchill Livingstone Elsevier.
- Schnyer, R., Lao, L., Hammerschlag, R., Wayne, P., Langevin, H. M., Napadow, V., Harris, R., Park, J., Milley, R., Cohen, M. & MacPherson, H. (2008) Society for Acupuncture Research: 2007 conference report: "The status and future of acupuncture research: 10 years post-NIH Consensus Conference", *Journal of alternative and complementary medicine*, 14, pp. 859-860. Doi: 10.1089/acm.2008.SAR-2.
- Schnyer, R. N. & Allen, J. J. (2002) Bridging the gap in complementary and alternative medicine research: manualization as a means of promoting standardization and flexibility of treatment in clinical trials of acupuncture, *Journal of alternative and complementary medicine*, 8, pp. 623.
- Shapiro, A. K. (1959) The placebo effect in the history of medical treatment: implications for psychiatry, *The American Journal Of Psychiatry*, 116, pp. 298-304.
- Sherman, K. J., Cherkin, D. C. & Hogeboom, C. J. (2001a) The diagnosis and treatment of patients with chronic low-back pain by traditional Chinese medical acupuncturists, *Journal of alternative and complementary medicine*, 7, pp. 641-650. <https://doi.org/10.1089/10755530152755199>
- Sherman, K. J., Hogeboom, C. J. & Cherkin, D. C. (2001b) How traditional Chinese medicine acupuncturists would diagnose and treat chronic low back pain: results of a survey of licensed acupuncturists in Washington State, *Complementary therapies in medicine*, 9, pp. 146-153. Doi: 10.1054/ctim.2001.0458,.
- Shi, X. m. & Zhang, M. c. (1998) A Chinese–English Dictionary of Acupuncture and Moxibustion. China: Huaxia Publishing House.
- Shodo, D. (2003) *Finding Effective Acupuncture Points*. Seattle: Eastland Press, Incorporated.
- Silverwood, V., Blagojevic-Bucknall, M., Jinks, C., Jordan, J. L., Protheroe, J. & Jordan, K. P. (2015) Current evidence on risk factors for knee osteoarthritis in older adults: a systematic review and meta-analysis, *Osteoarthritis And Cartilage*, 23, pp. 507-515. Doi: 10.1016/j.joca.2014.11.019.

- Singh, S. & Ernst, E. (2008) *Trick or Treatment? Alternative Medicine on Trial*. London, Toronto, Sydney, Auckland, Johannesburg: Bantam Press.
- Streitberger, K., Ezzo, J. & Schneider, A. (2006) Acupuncture for nausea and vomiting: an update of clinical and experimental studies, *Autonomic Neuroscience: Basic & Clinical*, 129, pp. 107-117.
- Streitberger, K. & Kleinhenz, J. (1998) Introducing a placebo needle into acupuncture research, *Lancet*, 352, pp. 364-365.
- Streitberger, K., Witte, S., Mansmann, U., Knauer, C., Krämer, J., Scharf, H.-P. & Victor, N. (2004) Efficacy and safety of acupuncture for chronic pain caused by gonarthrosis: a study protocol of an ongoing multi-centre randomised controlled clinical trial [ISRCTN27450856], *BMC Complementary And Alternative Medicine*, 4, pp. 6-6. <https://doi.org/10.1186/1472-6882-4-6>
- Sturgeon, D. (2018) *Chinese Text Project*. Available from: <https://ctext.org/> [Accessed 22 February 2018]
- Stux, G. & Birch, S. (2001). Proposed standards of acupuncture treatment for clinical studies. In: Hammerschlag, R. & Stux, G. (eds.) *Clinical acupuncture: scientific basis*. Berlin: Springer, pp. 171-185.
- Suarez-Almazor, M. E., Looney, C., Liu, Y., Cox, V., Pietz, K., Marcus, D. M. & Street, R. L., Jr. (2010) A randomized controlled trial of acupuncture for osteoarthritis of the knee: effects of patient-provider communication, *Arthritis Care & Research*, 62, pp. 1229-1236. Doi: 10.1002/acr.20225.
- Sun, G. J. 孙国杰 (2000) [The Study of Acupuncture] 针灸学. Beijing: The Peoples' Health Publishing House 人民卫生出版社
- Tan, H. y. & Feng, Y. l. 谭惠予 & 冯云亮 (2016) Wēn zhēnjiǔ shēn cì dú bí, nèi xī yǎn zhìliáo xī guānjié gǔ xìng guānjié yán de línchuáng liáoxiào fēnxī 温针灸深刺犊鼻、内膝眼治疗膝关节骨性关节炎的临床疗效分析 Clinical Analysis of Warm Needle Acupuncture in Treating Knee Joint Osteoarthritis, *Zhōngyī línchuáng yánjiū* 中医临床研究 (*Clinical Journal of Chinese Medicine*), Number 14, pp 23-24. [Chinese].
- Thabane, L., Ma, J., Chu, R., Cheng, J., Ismaila, A., Rios, L. P., Robson, R., Thabane, M., Giangregorio, L. & Goldsmith, C. H. (2010) A tutorial on pilot studies: the what, why and how, *BMC Medical Research Methodology*, 10, pp. 1-1. Doi: 10.1186/1471-2288-10-1.
- Tubach, F., Ravaud, P., Martin-Mola, E., Awada, H., Bellamy, N., Bombardier, C., Felson, D. T., Hajjaj-Hassouni, N., Hochberg, M., Logeart, I., Matucci-Cerinic, M., van de Laar, M., van der Heijde, D. & Dougados, M. (2012) Minimum clinically important improvement and patient acceptable symptom state in pain and function in rheumatoid arthritis, ankylosing spondylitis, chronic back pain, hand osteoarthritis, and hip and knee osteoarthritis: Results from a prospective multinational study, *Arthritis Care & Research*, 64, pp. 1699-1707. Doi: 10.1002/acr.21747.
- Tukmachi, E., Jubb, R., Dempsey, E. & Jones, P. (2004) The effect of acupuncture on the symptoms of knee osteoarthritis-an open randomised controlled study, *Acupuncture in Medicine*, 22. Doi: 10.1136/aim.22.1.14.
- Unschuld, P. U. (2003) *Huang Di nei jing su wen: nature, knowledge, imagery in an ancient Chinese medical text*. Berkeley: University of California Press.

- Unschuld, P. U. (2009) *What Is Medicine? Western and Eastern Approaches to Healing*. University of California Press.
- Unschuld, P. U. (2010) *Medicine in China: a history of ideas*. Berkeley: University of California Press.
- Urbaniak, G. C. & Plous, S. (2013) *Research Randomizer (Version 4.0) [Computer software]*. Available from: <https://www.randomizer.org> [Accessed December 2015]
- Vas, J., Méndez, C., Perea-Milla, E., Vega, E., Panadero, M. D., León, J. M., Borge, M. A., Gaspar, O., Sánchez-Rodríguez, F., Aguilar, I. & Jurado, R. (2004) Acupuncture as a complementary therapy to the pharmacological treatment of osteoarthritis of the knee: randomised controlled trial, *BMJ Clinical Research*, 329, pp. 1216-1216.
- Veith, I. (2016) *Huang di nei jing su wen.: the Yellow Emperor's classic of internal medicine : chapters 1-34*. Berkeley: University of California Press.
- Vickers, A. J., Cronin, A. M., Maschino, A. C., Lewith, G., Macpherson, H., Foster, N. E., Sherman, K. J., Witt, C. M., Linde, K. & for-the-Acupuncture-Trialists-Collaboration (2012) Acupuncture for Chronic Pain: Individual Patient Data Meta-analysis, *Archives of Internal Medicine*, pp. E1-10. Doi: 10.1001/archinternmed.2012.3654.
- Vickers, A. J., Maschino, A. C., Lewith, G., MacPherson, H., Sherman, K. J. & Witt, C. M. (2013) Responses to the Acupuncture Trialists' Collaboration individual patient data meta-analysis, *Acupuncture in medicine : journal of the British Medical Acupuncture Society*, 31, pp. 98-100. Doi: 10.1136/acupmed-2013-010312.
- Vickers, A. J., Vertosick, E. A., Lewith, G., MacPherson, H., Foster, N. E., Sherman, K. J., Irnich, D., Witt, C. M. & Linde, K. (2017) Acupuncture for Chronic Pain: Update of an Individual Patient Data Meta-Analysis, *The Journal Of Pain: Official Journal Of The American Pain Society*. Doi: 10.1016/j.jpain.2017.11.005.
- Vickers, J. (2016) *The Problem of Induction*: Stanford Encyclopedia of Philosophy. Available from: <<https://plato.stanford.edu/archives/spr2016/entries/induction-problem/>>. [Accessed 22 February 2018]
- Vincent, C. & Lewith, G. (1995) Placebo controls for acupuncture studies, *Journal of the Royal Society of Medicine*, 88, pp. 199-202.
- Vos, T., Flaxman, A. D., Naghavi, M. & et al. (2012) Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010, *The Lancet*, 380, pp. 2163-2196. Doi: [http://dx.doi.org/10.1016/S0140-6736\(12\)61729-2](http://dx.doi.org/10.1016/S0140-6736(12)61729-2).
- Walach, H., Falkenberg, T., Fønnebo, V., Lewith, G. & Jonas, W. B. (2006) Circular instead of hierarchical: methodological principles for the evaluation of complex interventions, *BMC Medical Research Methodology*, 6, pp. 1-9. Doi: 10.1186/1471-2288-6-29.
- Wang, F., Li, T., Yu, X., Deng, Y., Bo, J., Chen, F., Zhang, Q., Yu-tong, F., Fang, T. & Davies, C. (2007) *Scalp acupuncture therapy*. Beijing: People's Medical Publishing House.
- Wang, G. z. 王光宗 (2007) Wēnzhēnjiǔ zhìliáo xīgūānjié gǔxìng guānjiéyán 50 lì liáoxiào guānchá 温针灸治疗膝关节骨性关节炎50例疗效观察 [Observation of the therapeutic efficacy of warm needle acupuncture therapy for 50 cases of osteoarthritis of the knee], *Zhōngguó zhēnjiǔ xuéhuì línchuáng fēnhuì dì shíwǔ jiè quánáguó zhēnjiǔ xuéshù yántǎohuì* 中国针灸学会临床分会第十五届全国针灸学术研讨会 [The Clinical

- Committee of the Chinese Acupuncture Society's 15th National Academic Symposium on Acupuncture], pp 4. [Chinese].
- Wang, G. z. & Zhu, L. j. 王光宗 & 诸灵军 (2009) Wēn zhēnjiǔ zhìliáo lǎonián xī guānjié gǔ xìng guānjié yán 35 lì liáoxiào guānchá 温针灸治疗老年膝关节骨性关节炎35例疗效观察 [Observation of the therapeutic efficacy of warm needle acupuncture therapy for 35 cases of age-related osteoarthritis of the knee], *Zhōngguó zhēnjiǔ xuéhuì línchuáng fēnhuì dì shíqī jiè quánguó zhēnjiǔ línchuáng xuéshù yántǎohuì* 中国针灸学会临床分会第十七届全国针灸临床学术研讨会 [The Clinical Committee of the Chinese Acupuncture Society's 17th National Clinical Acupuncture Academic Symposium], pp 3. [Chinese].
- Wayne, P. M., Hammerschlag, R., Langevin, H. M., Napadow, V., Park, J. J. & Schnyer, R. N. (2009) Resolving paradoxes in acupuncture research: a roundtable discussion, *Journal of alternative and complementary medicine*, 15, pp. 1039-1044. Doi: 10.1089/acm.2009.0466.
- Wengraf, T. (2001) Qualitative research interviewing: biographic narrative and semi-structured methods. London ; Thousand Oaks, Calif: Sage.
- White, A. (2004) A cumulative review of the range and incidence of significant adverse events associated with acupuncture, *Acupuncture In Medicine: Journal Of The British Medical Acupuncture Society*, 22, pp. 122-133.
- White, A. & Cummings, M. (2015) Acupuncture for knee osteoarthritis: study by Hinman et al represents missed opportunities, *Acupuncture In Medicine: Journal Of The British Medical Acupuncture Society*, 33, pp. 84-86. Doi: 10.1136/acupmed-2014-010719.
- White, A., Cummings, M. & Filshie, J. (2008) *An Introduction to Western Medical Acupuncture*. Edinburgh, London, New York, Oxford, Philadelphia, St Louis, Sydney, Toronto: Churchill Livingstone Elsevier.
- White, A. & Editorial-Board (2009) Western medical acupuncture: a definition, *Acupuncture In Medicine: Journal Of The British Medical Acupuncture Society*, 27, pp. 33-35. Doi: 10.1136/aim.2008.000372.
- White, A., Foster, N. E., Cummings, M. & Barlas, P. (2007) Acupuncture treatment for chronic knee pain: a systematic review, *Rheumatology*, 46, pp. 384-390.
- White, A. R. & Ernst, E. (1999) A systematic review of randomized controlled trials of acupuncture for neck pain, *Rheumatology*, 38, pp. 143-147.
- White, I. R., Horton, N. J., Carpenter, J. & Pocock, S. J. (2011) Strategy for intention to treat analysis in randomised trials with missing outcome data, *British Medical Journal*, 342. Doi: 10.1136/bmj.d40.
- Williamson, L., Wyatt, M. R., Yein, K. & Melton, J. T. K. (2007) Severe knee osteoarthritis: a randomized controlled trial of acupuncture, physiotherapy (supervised exercise) and standard management for patients awaiting knee replacement, *Rheumatology*, 46, pp. 1445-1449. <http://rheumatology.oxfordjournals.org/content/46/9/1445.full.pdf>
- Winter, J. 2016. RE: Correspondence regarding Five-element acupuncture. Personal e-mail to Appleyard, I. 5th May.

- Witt, C., Brinkhaus, B., Jena, S., Linde, K., Streng, A., Wagenpfeil, S., Hummelsberger, J., Walther, H. U., Melchart, D. & Willich, S. N. (2005) Acupuncture in patients with osteoarthritis of the knee: a randomised trial, *Lancet*, 366, pp. 136-143.
- Witt, C. M. (2009) Efficacy, Effectiveness, Pragmatic Trials – Guidance on Terminology and the Advantages of Pragmatic Trials, *Forschende Komplementärmedizin / Research in Complementary Medicine*, 16, pp. 292-294. Doi: 10.1159/000234904.
- Witt, C. M. (2011) Clinical research on acupuncture - Concepts and guidance on efficacy and effectiveness research, *Chinese Journal of Integrative Medicine*, 17, pp. 166-172. Doi: 10.1007/s11655-011-0662-x.
- Witt, C. M., Aickin, M., Baca, T., Cherkin, D., Haan, M. N., Hammerschlag, R., Hao, J. J., Kaplan, G. A., Lao, L., McKay, T., Pierce, B., Riley, D., Ritenbaugh, C., Thorpe, K., Tunis, S., Weissberg, J. & Berman, B. M. (2012) Effectiveness Guidance Document (EGD) for acupuncture research - a consensus document for conducting trials, *BMC Complementary And Alternative Medicine*, 12, pp. 148-148. Doi: 10.1186/1472-6882-12-148.
- Witt, C. M., Jena, S., Brinkhaus, B., Liecker, B., Wegscheider, K. & Willich, S. N. (2006) Acupuncture in patients with osteoarthritis of the knee or hip: a randomized, controlled trial with an additional nonrandomized arm, *Arthritis and Rheumatism*, 54, pp. 3485-3493. Doi: 10.1002/art.22154.
- Witt, C. M., Pach, D., Brinkhaus, B., Wruck, K., Tag, B., Mank, S. & Willich, S. N. (2009) Safety of acupuncture: results of a prospective observational study with 229,230 patients and introduction of a medical information and consent form, *Forschende Komplementärmedizin*, 16, pp. 91-97. Doi: 10.1159/000209315.
- Wong Lit Wan, D., Wang, Y., Xue, C. C., Wang, L. P., Liang, F. R. & Zheng, Z. (2015) Local and distant acupuncture points stimulation for chronic musculoskeletal pain: A systematic review on the comparative effects, *European Journal of Pain*, pp. 1232-47. Doi: 10.1002/ejp.671.
- World-Medical-Association. (2013) *Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Subjects*. Available from: <https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/> [Accessed January 2014]
- Worsley, J. R. (1998) *Classical Five-Element Acupuncture, Vol. 3: The Five Elements and the Officials*. Worsley, Inc.
- Worsley, J. R. (2012) *Worsley Five-Element Acupuncture, Vol II Traditional Diagnosis*. Miami: Worsley, Inc.
- Wu, H. m. 吴红梅 (2015) Wēn zhēnjiǔ zhìliáo xī guānjié gǔ guānjié yán de jìnqī xiàoguǒ yánjiū 温针灸治疗膝关节骨关节炎的近期效果研究 [Short - term effect of warm needle acupuncture on knee osteoarthritis], *Shijie zuìxīn yīxué xīnxī wénzhāi* 世界最新医学信息文摘 (*World Latest Medicine Information*), Number 34, pp 104. [Chinese].
- Wu, X. & Sheng, C. R. 吴旭 & 盛灿若 (2002) *The Study of Practical Acupuncture and Moxibustion 实用针灸学*. Jiangsu Education Publishing House 江苏教育出版社 [Chinese].
- Wu, X. f. 吴秀芬 (1999) *Traditional Chinese Diagnostics 中医诊断学*. Beijing: People's Medical Publishing House 人民卫生出版社 [Chinese].

- Xia, Q. f. 夏秋芳 (2011) Fùzhēn jiéhé júbù wēnzhēnjiǔ zhìliáo xīgǔxìng guānjiéyán liáoxiào guānchá 腹针结合局部温针灸治疗膝骨性关节炎疗效观察 [Observation of the therapeutic efficacy of abdominal acupuncture combined with local warmer needle acupuncture therapy for osteoarthritis of the knee], *Zhēnjiǔ línchuáng zázhi* 针灸临床杂志 (*Journal of Clinical Acupuncture and Moxibustion*), Number 03, pp 16-17. [Chinese].
- Xiong, G.-p., Huang, Q. & Dong, Y.-f. (2011) Duǎncì pèihé wēnzhēnjiǔ zhìliáo xīgūānjié gǔxìng guānjiéyán 30 lì 短刺配合温针灸治疗膝关节骨性关节炎30例 [Thirty cases of knee osteoarthritis treated by acupuncture and warming needle moxibustion], *Zhōngguó zhēnjiǔ* 中国针灸 *Chinese Acupuncture & Moxibustion*, Number 6, pp 551-552.
- Xu, G. y. 许贵英 (2016) Wēn zhēnjiǔ zhìliáo hán shī xíng xī guānjié gǔ xìng guānjié yán de liáoxiào 温针灸治疗寒湿型膝关节骨性关节炎的疗效 Efficacy of Warm Needle Acupuncture in the Treatment of Osteoarthritis of the Knee, *Zhōngguó zhōng yīyào xiàndài yuǎnchéng jiàoyù* 中国中医药现代远程教育 (*Chinese Medicine Modern Distance Education of China*), Number 20, pp 107-108. [Chinese].
- Xu, H. Z., Ni, Y. T., Liu, Y. G. & Wu, Z. C. (1988) Acupuncture treatment of common diseases: Based upon differentiation of syndromes. Beijing China: The People's medical publishing house.
- Yan, X. l., Zhang, W. j. & Dong, B. q. 闫晓黎, 张文静 & 董宝强 (2013) Wēnzhēnjiǔ zhìliáo xīgǔxìng guānjiéyán suíjī píngxíng duìzhào yánjiū 温针灸治疗膝骨性关节炎随机平行对照研究 (Warm acupuncture treatment of knee osteoarthritis randomised controlled study), *Shíyòng zhōngyī nèikē zázhi* 实用中医内科杂志 (*Journal of Practical Traditional Internal Chinese Medicine*), Number 08, pp 128-130. [Chinese].
- Yang, D. 杨冬 (2014) Wēnzhēnjiǔ hé pǔtōng zhēncì zhìliáo yánjiū zhìliáo tuìxíngxìng xīgǔ guānjiéyán (yángxūhán níngxíng) línchuáng yánjiū 温针灸和普通针刺治疗研究 治疗退行性膝骨关节炎(阳虚寒凝型)临床研究 [Clinical research of warm needle acupuncture and common acupuncture needle insertion therapy for degenerative knee osteoarthritis (Yang deficiency congealing cold type)], *Dàjiā jiànkāng (xuéshù bǎn)* 大家健康(学术版) (*Health for All*), Number 02, pp 47-48. [Chinese].
- Yang, J. z. 杨继洲 (1601) 针灸大成 *Great Compendium of Acupuncture and Moxibustion*, 明清中医临证小丛书 [The Qing and Ming dynasties clinical Chinese medicine series], Published 中国中医药出版社 (China Press of Traditional Chinese Medicine), 北京 Beijing. [Chinese].
- Yang, K. y. & Qu, X. x. 杨开洋 & 屈箫箫 (2015) Wēn zhēnjiǔ zhìliáo xī guānjié gǔ xìng guānjié yán de línchuáng guānchá 温针灸治疗膝关节骨性关节炎的临床观察 [Clinical Observation on Treatment of Osteoarthritis of Knee Joint with Warm Needle Acupuncture], *Shǎnxī zhōngyī* 陕西中医 (*Shaanxi Journal of Traditional Chinese Medicine*), Number 11, pp 1520-1521. [Chinese].
- Yang, X. s. 杨学双 (2016) Dānchún zhēn cì yǔ wēn zhēnjiǔ zhìliáo lǎonián xī gǔ guānjié bìng xiàoguǒ de línchuáng bǐjiào 单纯针刺与温针灸治疗老年膝骨关节病效果的临床比较 [Clinical comparison of needle acupuncture and warm needle technique in the treatment of knee osteoarthritis in elderly patients], *Zhōngguó yīyào zhǐnán* 中国医药指南 (*Guide of China Medicine*), Number 35, pp 191-192. [Chinese].

- Yao, Z. f., Huang, X. q. & Zhang, W. 姚志芳, 黄晓卿 & 张炜 (2003) Wēnzhēnjiǔ yǔ dānchún zhēncì duì xīguānjié gǔxìng guānjiéyán liáoxiào bǐjiào guānchá 温针灸与单纯针刺对膝关节骨性关节炎疗效比较观察 [An observation of the comparative clinical efficacy of warm needle acupuncture and simple needle insertion for osteoarthritis of the knee], *Zhēnjiǔ línchuáng zázhi* 针灸临床杂志 (*Journal of Clinical Acupuncture and Moxibustion*), Number 07, pp 34-35. [Chinese].
- Yu, J. 余健 (2008) Wēnzhēnjiǔ zhìliáo tuìxíngxìng xīguānjié bìng 62 lì de línchuáng guānchá 温针灸治疗退行性膝关节病 62 例的临床观察 [Clinical Observation of 62 cases of degenerative knee joint disease treated with warm needle acupuncture], *Quánguó dì 16 jiè zhēnjiǔ línchuáng xuéshù yántǎohuì, quánguó dì 11 jiè ěr xué zhēnzhì xuéshù yántǎohuì, dāngdài línchuángzhì yànlùntán jì zhōng xībù shíshěngqū xuéshù yántǎohuì* 全国第16 届针灸临床学术研讨会、全国第11届耳穴诊治学术研讨会、当代临床治验论坛暨中西部十省区学术研讨会 (*The 16th National Acupuncture and Moxibustion Academic Deliberation Conference*), pp 3. [Chinese].
- Yu, J. 余健 (2009) Wēnzhēnjiǔ zhìliáo xīguānjié gǔxìng guānjiéyán 32 lì guānchá 温针灸治疗膝关节骨性关节炎32例观察 [A clinical observation of warm needle acupuncture therapy for 62 cases of degenerative knee joint disease], *Shíyòng zhōng yīyào zázhi* 实用中医药杂志 (*Journal of Practical Traditional Chinese Medicine*), Number 02, pp 93. [Chinese].
- Yue, R. 岳蓉 (2010) Tòucì wēnzhēnjiǔ zhìliáo xīgǔ guānjiéyán liáoxiào guānchá 透刺温针灸治疗膝骨关节炎疗效观察 [Observation of the therapeutic efficacy of through needle warm acupuncture for osteoarthritis of the knee], *Línchuáng hé lǐ yòngyào zázhi* 临床合理用药杂志 (*Chinese Journal of Clinical Rational Drug Use*), Number 15, pp 65. [Chinese].
- Zhang, C. S., Tan, H. Y., Zhang, G. S., Zhang, A. L., Xue, C. C. & Xie, Y. M. (2015) Placebo Devices as Effective Control Methods in Acupuncture Clinical Trials: A Systematic Review, *Plos One*, 10, pp. e0140825. Doi: 10.1371/journal.pone.0140825.
- Zhang, D. (2002) *Key Concepts of Chinese Philosophy*. Trs Ryden, E. New Haven and London, Beijing: Yale University and Foreign Languages Press
- Zhang, G. c. 张广才 (2015) Wēn zhēnjiǔ yǔ zhēn cì zhìliáo lǎonián xìng xī gǔ guānjié bìng de línchuáng liáoxiào guānchá 温针灸与针刺治疗老年性膝骨关节病的临床疗效观察 [Clinical Observation on Treatment of Old Aged Knee Osteoarthrosis by Warm Needle Acupuncture and Needle Acupuncture], *Dàjiā jiànkāng (xuéshù bǎn)* 大家健康 (学术版) (*Everybody Healthy [Academic Edition]*), Number 11, pp 24-25. [Chinese].
- Zhang, J. 张菊 (2013a) Wēnzhēnjiǔ zhìliáo xīgǔ guānjiéyán 90 lì 温针灸治疗膝骨关节炎 90例 [The treatment of 90 cases of osteoarthritis of the knee with warm needle acupuncture], *Zhōngguó zhōngyīyào xiàndài yuǎnchéng jiàoyù* 中国中医药现代远程教育 (*Chinese Medicine Modern Distance Education of China*), Number 22, pp 77. [Chinese].
- Zhang, J. f., Wu, Y. c. & Li, S. s. 张峻峰, 吴耀持 & 李石胜 (2009) Tòucì wēnzhēnjiǔ zhìliáo xīgǔ guānjiéyán liáoxiào guānchá 透刺温针灸治疗膝骨关节炎疗效观察 (Observations on the efficacy of Point through Point acupuncture and moxibustion in treating genua osteoarthritis), *Shànghǎi zhēnjiǔ zázhi* 上海针灸杂志 *Shanghai Journal of Acupuncture and Moxibustion*, Number 12, pp 722-723. [Chinese].

- Zhang, Q., Yue, J., Golianu, B., Sun, Z. & Lu, Y. (2017) Updated systematic review and meta-analysis of acupuncture for chronic knee pain, *Acupuncture In Medicine: Journal Of The British Medical Acupuncture Society*, 35, pp. 392-403. Doi: 10.1136/acupmed-2016-011306.
- Zhang, W., Robertson, J., Jones, A. C., Dieppe, P. A. & Doherty, M. (2008) The placebo effect and its determinants in osteoarthritis: meta-analysis of randomised controlled trials, *Annals of the Rheumatic Diseases*, 67, pp. 1716-1723. Doi: 10.1136/ard.2008.092015.
- Zhang, Z. J., Wang, X. M. & McAlonan, G. M. (2012) Neural acupuncture unit: a new concept for interpreting effects and mechanisms of acupuncture, *Evid Based Complement Alternat Med*, 2012, pp. 429412. Doi: 10.1155/2012/429412.
- Zhang, Z. p. 张芝平 (2013b) Qiǎn xī wēn zhēnjiǔ zhìliáo lǎonián xī gǔ guānjié bìng 43 lì línchuáng liáoxiào 浅析温针灸治疗老年膝骨关节病43例临床疗效 (Analysis of the clinical effect on treating 43 cases of senile knee osteoarthropathy by warm acupuncture), *Zhōngyī línchuáng yánjiū* 中医临床研究 (Clinical Journal of Chinese Medicine), Number 21, pp 37. [Chinese].
- Zhao, J. & Wang, Y.-p. (2007) Chinese medicine study guide. Acupuncture and moxibustion. Beijing: People's Medical Publishing House.
- Zhao, J. S. (ed.) (2002). *Chinese Acupuncture and Moxibustion*: Publishing House of Shanghai University of Traditional Chinese Medicine.
- Zhao, M. 赵敏 (2014) Wēn zhēnjiǔ zhìliáo xī guānjié gǔ xìng guānjié yán 35 lì línchuáng guānchá 温针灸治疗膝关节骨性关节炎35例临床观察 [Clinical observation of the treatments of 35 cases of osteoarthritis of the knee by warm needle acupuncture], *Shǎnxī zhōngyī* 陕西中医 *Shaanxi Journal of Traditional Chinese Medicine*, Number 03, pp 359-360. [Chinese].
- Zhao, X. t., Lu, X. d., Li, Z. w. & Lin, S. m. 赵学田, 卢向东, 李兆文 & 林石明 (2008) Wēnzhēn cǐfǎ zhìliáo xīguānjié gǔxìng guānjiéyán 50 lì 温针刺法治疗膝关节骨性关节炎50例 [50 cases of osteoarthritis of the knee treated with warm needle acupuncture method], *Zhōngyīyào tōngbào* 中医药通报 *Traditional Chinese Medicine Journal*, Number 04, pp 49-50. [Chinese].
- Zheng, X. y. 郑筱萸 (2002) Guiding principles of clinical research of new medicine in Chinese medicine 中药新药临床研究指导原则. Beijing: China Medical Science Press 中国医药科技出版社. [Chinese].
- Zhu, L., Jin, Z., Liao, L. q., Xu, L. k. & Li, Y. 朱兰, 金柱, 廖丽群, 徐立坤 & 黎云 (2011) Wēnzhēnjiǔ liáofǎ zhìliáo zǎoqī lǎonián xīguānjié gǔxìng guānjiéyán 40 lì línchuáng yánjiū 温针灸疗法治疗早期老年膝关节骨性关节炎40例临床研究 [A clinical observation of the treatment of liver and kidney depletion type osteoarthritis of the knee by warm needle acupuncture], *Zhōngguó shèqū yīshī (yīxué zhuānyè)* 中国社区医师(医学专业) *Chinese Community Doctors (Medical Professionals)*, Number 21, pp 190-191. [Chinese].