

**THE EFFECTIVENESS OF PHYSIOTHERAPY FOLLOWING  
DISCHARGE FROM HOSPITAL AFTER PRIMARY TOTAL  
KNEE ARTHROPLASTY FOR OSTEOARTHRITIS.**

by

CATHERINE JANE MINNS LOWE

A thesis submitted to the University of Birmingham  
for the degree of Ph.D

Department of Primary Care and General Practice  
School of Health and Population Sciences  
The University of Birmingham  
2009

UNIVERSITY OF  
BIRMINGHAM

**University of Birmingham Research Archive**

**e-theses repository**

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

## **ABSTRACT.**

This thesis evaluates and explores the effectiveness of post discharge physiotherapy exercise following total knee arthroplasty (TKA) for osteoarthritis in three ways.

1. A systematic review evaluated the effectiveness of post discharge physiotherapy exercise on function, walking, range of motion, quality of life and muscle strength, for patients following elective primary TKA. Functional physiotherapy exercise interventions following discharge resulted in short term, but not long term, benefit. Effect sizes were small to moderate for function (0.33). Weighted mean differences were small to moderate for motion (2.9) and small for quality of life (1.66).

2. A randomised clinical trial compared the effectiveness of a post discharge physiotherapy intervention in improving patient function versus usual physiotherapy for patients undergoing primary TKA. No significant statistical differences were observed between the two groups for all outcomes. This early trial was underpowered and impacted upon by some important factors which could potentially have masked any treatment trends occurring in the home visit group.

3. Since blinding procedures are often assumed to indicate trial quality, the feasibility of achieving blind outcome assessment in a pragmatic physiotherapy rehabilitation trial involving older people was explored. Reasons for unblinding were explored and successful blinding rates of 81-91% were achieved.

## **DEDICATION.**

To Tim and Iona Lowe, Rev. John Minns  
and to R.M.L., A.M.L. and A.W.M.L.

*Lux et Umbra vicissim, sed semper Amor*



## **ACKNOWLEDGEMENTS.**

I would like to acknowledge and thank the following people and organisations for their assistance during the preparation of this thesis:

My supervisors, Professor Cath Sackley and Dr Karen Barker, for their continuing support, assistance, advice and commitment to both the completion of this thesis and the development of my research career. I am very grateful for all your help and time.

Michelle Wilson, Kathleen Reilly and all other members of the Physiotherapy Research Unit at the Nuffield Orthopaedic Centre NHS Trust in Oxford who assisted with recruitment, data collection, and providing the intervention in the randomised clinical trial component of this thesis.

The orthopaedic surgeons and staff in the Pre Operative Assessment Clinic at the Nuffield Orthopaedic Centre NHS Trust in Oxford without whose cooperation and help the research would not have been possible.

All the patients who participated in the randomised clinical trial. Thank you so much for your time and assistance. And to Mr Trevor Scholey, who kindly agreed to participate in the development of the trial procedures as a lay member of the trial steering committee and whose input into the participant information sheets and development of the intervention was much appreciated.

Professor Mike Clarke and the students on the “Systematic Reviews” Module, May 2005, University of Oxford Department for Continuing Education for their comments on the design of the systematic review.

Robert Bourne, David Beverland, P. Codine, Helen Frost, Patricia Humphreys, John Kramer and Brian Mockford for providing additional data for the systematic review included in this thesis.

The National Institute of Health Research for awarding an Allied Health Professional Researcher Development Award which funded me for the duration of this Ph.D.

To Oxfordshire PCT for providing additional funding to provide the physiotherapy interventions in the randomised controlled trial.

To Paul Cooper at the Nuffield Orthopaedic Hospital NHS Trust for supplying radiographic images of an osteoarthritic knee for use in this thesis.

To Tim and Iona. To Tim for encouraging, supporting and proof reading throughout. And to Iona, whose arrival both delayed the completion of this Ph.D. and placed her own perspective upon it.

# CONTENTS.

	<b>Page Number</b>
<b>LIST OF PUBLICATIONS AND ABSTRACTS DERIVED FROM THE RESEARCH.....</b>	<b>11</b>
<b>PUBLICATIONS.....</b>	<b>11</b>
<b>1 CHAPTER ONE. INTRODUCTION.....</b>	<b>13</b>
1.1 Introduction To Thesis.....	13
1.2 Developments Altering The Thesis.....	14
1.3 Aims And Hypotheses Of This Thesis.....	15
1.3.1 Part One. The Systematic Review.....	15
1.3.2 Background.....	15
1.3.2.1 Aim Of The Systematic Review.....	15
1.3.2.2 Objective For The Systematic Review.....	15
1.3.3 Part Two. The Randomised Clinical Trial.....	16
1.3.3.1 Background.....	16
1.3.3.2 Aim Of Trial.....	16
1.3.3.3 Hypotheses For The Trial.....	17
1.3.4 Part Three. The Evaluation Of Blind Outcome Assessment.....	18
1.3.4.1 Background.....	18
1.3.4.2 Aim Of Study.....	19
1.3.4.3 Objectives Of The Study.....	19
1.4 Overview Of Chapters.....	19
<b>2 CHAPTER TWO. BACKGROUND LITERATURE REVIEW.....</b>	<b>21</b>
2.1 Chapter Overview.....	21
2.2 Definitions Of Osteoarthritis.....	21
2.3 Signs And Symptoms Of Osteoarthritis Of The Knee.....	23
2.4 Aetiology Of Osteoarthritis Of The Knee.....	25
2.5 The Prevalence Of Osteoarthritis Of The Knee And The Rate Of Primary Total Knee Joint Replacement In The Uk.....	30
2.6 Current Concepts For Treatment Of Osteoarthritis Of The Knee.....	32
2.7 Joint Arthroplasty.....	33
2.8 Outcome After Joint Arthroplasty.....	36
2.9 Rehabilitation.....	38
2.9 Chapter Summary.....	43

<b>3</b>	<b>CHAPTER THREE. EFFECTIVENESS OF PHYSIOTHERAPY EXERCISE FOLLOWING KNEE ARTHROPLASTY FOR OSTEOARTHRITIS: A SYSTEMATIC REVIEW OF RANDOMISED CONTROLLED TRIALS.....</b>	<b>44</b>
3.1	Chapter Overview.....	44
3.2	Systematic Review: Introduction And Background.....	44
3.2.1	Section Overview.....	44
3.2.2	Definition.....	45
3.2.3	Background.....	45
3.2.4	Section Summary.....	51
3.3	Protocol.....	51
3.3.1	Section Overview.....	51
3.3.2	Step 1. Formulate Question For Review.....	52
3.3.3	Step 2. Definition Of Inclusion And Exclusion Criteria.....	53
3.3.4	Step 3. Location Of Studies.....	54
3.3.5	Step 4. Selection Of Studies.....	55
3.3.6	Step 5. The Assessment Of Study Quality.....	56
3.3.7	Step 6. Data Extraction.....	57
3.3.8	Step 7. Analysis And Presentation Of Results.....	58
3.3.9	Step 8. Interpretation Of Results.....	62
3.4	Results.....	62
3.4.1	Section Overview.....	62
3.4.2	Location Of Studies.....	63
3.4.3	Selection Of Studies.....	65
3.4.4	Pilot Study: The Assessment Of Study Quality.....	66
3.4.5	Main Review: The Assessment Of Study Quality.....	67
3.4.6	Masking In Knee Systematic Review:.....	69
3.4.7	Reviewer Agreement In Main Review:.....	69
3.4.8	Study Characteristics.....	69
3.4.9	Summary Of The Interventions And Comparisons.....	71
3.5	Data Synthesis.....	73
3.5.1	Measures Of Function (5 Trials, 494 Participants).....	73
3.5.2	Walking (3 Trials, 284 Participants).....	75
3.5.3	Range Of Joint Motion (5 Trials, 537 Participants).....	76
3.5.4	Quality Of Life (3 Trials, 387 Participants ).....	77
3.5.5	Muscle Strength.....	79
3.5.6	Heterogeneity.....	79
3.6	Discussion Of The Systematic Review.....	81

3.6.1	Section Overview.....	81
3.6.2	Summary Of Key Findings.....	81
3.6.3	Strengths And Weaknesses Of Review Procedures.....	81
3.6.3.1	Location Of Trials.....	81
3.6.3.2	Data Extraction.....	82
3.6.3.3	Trial Quality.....	84
3.6.3.4	Data Analysis.....	85
3.6.3.5	Clinical Implications.....	86
3.7	Chapter Summary.....	87
<b>4</b>	<b>CHAPTER FOUR. COMPARISON OF POST DISCHARGE PHYSIOTHERAPY VERSUS USUAL CARE FOLLOWING PRIMARY TOTAL KNEE ARTHROPLASTY FOR OSTEOARTHRITIS: A RANDOMISED CLINICAL TRIAL.....</b>	<b>89</b>
4.1	Chapter Overview.....	89
4.2	Purpose Of Trial.....	89
4.3	Methodological Approach.....	90
4.3.1	Section Overview.....	90
4.3.2	Definitions.....	90
4.3.3	The Value Of Randomised Clinical Trials.....	91
4.3.4	Limitations Of Randomised Clinical Trials.....	95
4.3.4.1	Experimentation May Be Inappropriate.....	97
4.3.4.2	Experimentation May Be Inadequate.....	100
4.3.5	How Much Pragmatism Is Permissible?.....	103
4.3.6	Section Summary.....	104
4.4	The Development Of The Trial Intervention.....	105
4.4.1	Section Overview.....	105
4.4.2	The Theoretical Model/Basis Of The Intervention.....	105
4.4.3	The Specific Actions And Activities Of The Intervention.....	113
4.4.3.1	Repetitive Task Training.....	122
4.4.3.2	Individual Concerns.....	125
4.5	Trial Procedures.....	125
4.5.1	Ethics.....	125
4.5.2	Design.....	125
4.5.3	Participants And Setting.....	126
4.5.4	Sample Size.....	126
4.5.5	Blinding.....	127
4.5.6	Outcome Measures.....	128
4.5.6.1	Primary Measure: The Oxford Knee Score (Oks).....	128
4.5.6.2	Secondary Measures:.....	130
4.5.6.3	The Knee Injury And Osteoarthritis Outcome Score (Koos.).....	130
4.5.6.4	The Leg Extensor Press (Lep).....	132
4.5.6.5	The Timed Walk Test.....	133
4.5.6.6	The Timed Sit To Stand (Sts) Test.....	135

4.5.6.7	Additional Measures.....	136
4.5.7	Procedure And Intervention:.....	138
4.5.7.1	Stage I. Recruitment Of Patients.....	138
4.5.7.2	Stage Ii. Randomisation And Allocation Concealment.....	141
4.5.7.3	Stage Iii. Surgery.....	142
4.5.7.4	Stage Iv. The Intervention.....	142
4.5.7.5	Stage V. Follow Up.....	143
4.5.8	Data Analysis.....	143
4.5.8.1	Missing Or Multiple Data.....	143
4.5.8.2	Data Analysis.....	144
4.6	Results.....	146
4.6.1	Flow Of Participants Through The Trial.....	146
4.6.2	Comorbidities, Mortality And Complication Rates Reported During The Trial.....	147
4.6.2.1	Comorbidity And Mortality.....	147
4.6.2.2	Complications Reported By Participants At Three Month And Twelve Month Follow Up Assessments.....	148
4.6.3	Further Interventions And Hospital Admissions Following Surgery.....	152
4.6.4	Further Joint Replacement Surgery.....	153
4.6.5	Baseline Characteristics.....	153
4.6.5.1	Descriptive Statistics For The Primary Outcome At Baseline.....	153
4.6.6	Baseline Characteristics For The Two Trial Groups.....	157
4.6.7	Baseline Outcomes For Each Group.....	159
4.7	Postoperative Follow Up Data At Three, Six And 12 Month Timepoints.....	161
4.7.1	Assessment Of Normality Within The Oxford Knee Score Primary Outcome Follow Up Data.....	161
4.7.2	Distribution Of Secondary Outcome Follow Up Data.....	162
4.7.3	Descriptive Analysis Of Primary And Secondary Outcome Follow Up Data.....	162
4.7.4	Within Group Comparisons Of Primary And Secondary Outcome Follow Up Data.....	164
4.7.5	Between Group Comparisons Of Secondary Outcome Follow Up Data.....	167
4.7.6	One-Way Repeated Measures Analysis Of Variance.....	168
4.7.7	Range Of Knee Joint Motion.....	175
4.7.8	The Physiotherapy Intervention.....	176
4.7.9	Additional Physiotherapy Provided To Participants During The Trial.....	179
4.7.10	Sample Size.....	181
4.7.10.1	Power.....	181
4.7.10.2	Sample Size.....	181
4.8	Discussion.....	182
4.8.1	Section Overview.....	182
4.8.2	Summary Of Main Findings.....	182
4.8.2.1	Between Group Differences.....	182
4.8.2.2	Within Group Differences.....	183
4.8.2.3	Repeated Measures.....	183
4.8.3	Strengths And Limitations Of This Research.....	184
4.8.3.1	Trial Procedures.....	184

4.8.3.2	Randomisation And Allocation Concealment .....	184
4.8.3.3	Sample Size .....	186
4.8.3.4	Complications Rates .....	187
4.8.3.5	Retention/Follow Up .....	190
4.8.3.6	The Intervention .....	190
4.8.4	Use Of Outcome Measures .....	192
4.8.4.1	Oxford Knee Score .....	192
4.8.4.2	Koos.....	195
4.8.4.3	Leg Extensor Press .....	196
4.8.4.4	Timed Walk .....	197
4.8.4.5	Timed Sit To Stand.....	198
4.8.5	How Does This Trial Compare To Previous Trials?.....	199
4.8.6	Implications For Clinical Practice .....	201
4.8.7	Future Research .....	203
4.9	Chapter Summary.....	205
<b>5</b>	<b>CHAPTER FIVE. BLIND OUTCOME ASSESSMENT IN A PRAGMATIC PHYSIOTHERAPY REHABILITATION TRIAL.....</b>	<b>206</b>
5.1	Chapter Overview.....	206
5.2	Introduction .....	206
5.3	Definitions And Terminology .....	207
5.4	Background .....	208
5.4.1	Effects Of Blinding.....	211
5.4.2	Feasibility Of Achieving Blind Outcome Assessment .....	212
5.4.3	Measurement And Analysis Of Blind Outcome Assessment .....	214
5.4.4	Section Summary.....	219
5.5	Purpose Of Study.....	219
5.6	Method.....	219
5.6.1	Section Overview .....	219
5.6.2	Protocol To Minimise Instances Of Unblinding .....	220
5.6.3	Recording Instances Of Unblinding .....	221
5.7	Data Analysis.....	222
5.8	Results .....	222
5.8.1	Questionnaire Findings.....	222
5.8.2	Diary Data .....	225
5.9	Discussion .....	227
5.9.1	Section Overview .....	227
5.9.2	Summary Of Key Findings.....	227
5.9.3	Strengths And Limitations Of The Study's Procedures.....	227

5.9.4	Implications Of The Study.....	231
5.9.5	Future Directions.....	234
5.10	Chapter Summary.....	236
<b>6</b>	<b>CHAPTER SIX. A SUMMARY OF THE MAIN FINDINGS, IMPLICATIONS AND FUTURE RESEARCH SUGGESTIONS IN THIS THESIS.....</b>	<b>237</b>
6.1	Chapter Overview.....	237
6.2	Systematic Review.....	237
6.2.1	Summary Of Main Findings.....	237
6.2.2	Implications Of The Research.....	238
6.2.3	Future Research Suggestions.....	238
6.3	Randomised Clinical Trial.....	239
6.3.1	Summary Of Main Findings.....	239
6.3.2	Implications Of The Research.....	239
6.3.3	Future Research Suggestions.....	240
6.4	Blind Outcome Assessment In A Pragmatic Physiotherapy Rehabilitation Trial.....	241
6.4.1	Summary Of Main Findings.....	241
6.4.2	Implications Of The Research.....	241
6.4.3	Future Research Suggestions.....	242
6.5	Conclusions Of This Thesis.....	243
6.5.1	Systematic Review.....	243
6.5.2	The Randomised Clinical Trial.....	243
6.5.3	Blind Outcome Assessment.....	244
6.6	Overall Thesis.....	244
	<b>REFERENCES.....</b>	<b>245</b>
	<b>APPENDIX I. Permission to reproduce Figure 1.....</b>	<b>270</b>
	<b>APPENDIX II. LANGUAGE BIAS: ITS IMPLICATIONS FOR A PROPOSED SYSTEMATIC REVIEW.....</b>	<b>271</b>
	<b>APPENDIX III. 22 Item Checklist. (based on the CONSORT<sup>1</sup> statement and CASP<sup>2</sup> guidelines.).....</b>	<b>277</b>
	<b>APPENDIX IV. ASSESSING THE QUALITY OF RANDOMISED CLINICAL TRIAL REPORTS FOR A PROPOSED REVIEW: IS BLINDING NECESSARY?.....</b>	<b>279</b>
	<b>APPENDIX V. Blinding Monitoring Form.....</b>	<b>286</b>
	<b>APPENDIX VI.....</b>	<b>288</b>
	<b>APPENDIX VII. LITERATURE SEARCH STRATEGIES.....</b>	<b>293</b>
	<b>APPENDIX VIII. REASONS FOR THE SELECTION OF PAPERS INCLUDED IN /EXCLUDED FROM THE REVIEW.....</b>	<b>301</b>
	<b>APPENDIX IX. STATISTICAL TESTS FOR REVIEWER AGREEMENT IN THE SYSTEMATIC REVIEW.....</b>	<b>304</b>
	<b>APPENDIX X. TRIAL EXERCISE BOOKLET.....</b>	<b>308</b>
	<b>APPENDIX XI. COPIES OF ETHICS AND TRUST MANAGEMENT APPROVAL AND INDEMNITY LETTERS.....</b>	<b>309</b>
	<b>APPENDIX XII. DISC OF QUESTIONNAIRES USED IN THE TRIAL.....</b>	<b>310</b>
	<b>APPENDIX IV. HOME VISIT CASE REPORT FORMS.....</b>	<b>315</b>



<b>APPENDIX XV. DISC OF STATISTICAL TESTS USED IN THE TRIAL.....</b>	<b>316</b>
<b>APPENDIX XVI. COPIES OF KNEE AND HIP SYTEMATIC REVIEW PUBLICATIONS.....</b>	<b>317</b>

## LIST OF ILLUSTRATIONS.

### Page Number

Illustration 1. Drawing of an osteoarthritic knee .....	26
Illustration 2. Radiograph showing an anterior-posterior view of an osteoarthritic knee joint with medial joint space narrowing and subchondral bone thickening. ....	26
Illustration 3. Radiograph showing a lateral view of an osteoarthritic knee joint .....	27
Illustration 4. Radiograph showing an anterior posterior view of a total knee joint arthroplasty.....	34
Illustration 5. Radiograph showing lateral view of a total knee arthroplasty. ....	35
Illustration 6. Photograph of Exercise 1. ....	116
Illustration 7. Photograph of Exercise 2. ....	117
Illustration 8. Photograph of Exercise 3. ....	118
Illustration 9. Photograph of Exercise 4. ....	118
Illustration 10. Photograph of Exercise 5. ....	118
Illustration 11. Photograph of Exercise 6. ....	119
Illustration 12. Photograph of Exercise 7. ....	120
Illustration 13. Photograph of Exercise 8. ....	120
Illustration 14. Photographs of Exercise 9.....	121
Illustration 15. Photograph of getting out of a car.....	122
Illustration 16. Photograph of getting up from a chair at a table. ....	123
Illustration 17. Photograph of walking outside.....	123
Illustration 18. Photograph of descending stairs.....	124
Illustration 19. Photograph showing the leg extensor press in use. ....	140

## LIST OF TABLES.

	<b>Page Number</b>
Table 1. Summary of the advantages and limitations of systematic reviews for Physiotherapy (Summarised using Mulrow, 1994; Greenhalgh, 1997; Hopayian, 2001; Main, 2003; Pettigrew, 2003).	49
Table 2. Search Strategy for Systematic Review	64
Table 3. Studies Excluded from the knee systematic review.	66
Table 4. Quality component checklist and quality evaluation of trials included in the review (n=6).	68
Table 5. Study Characteristics of the Trials Evaluated in the Knee Systematic Review (n=6). TKA denotes total knee arthroplasty.	70
Table 6. Summary of trial interventions and comparisons included in the knee replacement trials (n=6).	72
Table 7. Heterogeneity $\chi^2$ and $I^2$ Test Results for Function, Walking, Joint Range of Motion and Quality of Life.	80
Table 8. Post operative major complications related to knee replacement surgery for 0-3 months (n=92) and 4-12 months following surgery (n=93).	149
Table 9. Table to show treatment arm allocation for post operative major complications related to knee replacement surgery for 0-3 months (n=92) and 4-12 months following surgery (n=93).	150
Table 10. Table to describe the data obtained for the primary outcome at baseline (n=105).	154
Table 11. Baseline characteristics of the Home Visit and the Usual Physiotherapy Groups.	158
Table 12. Baseline Outcome scores for the Home Visit and the Usual Physiotherapy Groups.	160
Table 13. Table to describe the data obtained for the primary outcome at all timepoints (n=105).	161
Table 14. Kolmogorov-Smirnov statistical test results for the Oxford Knee Score at all timepoints.	162
Table 15. Outcome scores for the Home Visit (HV) and the Usual Physiotherapy Care (UPC) Groups at baseline, 3, 6, and 12 months.	163
Table 16. Table presenting the comparisons of scores within groups at baseline, 3, 6, and 12/12 timepoints.	165
Table 17. Table presenting the primary outcome Oxford knee score median change scores within the two groups for all time points.	165
Table 18. Table presenting the secondary outcome mean / median change scores within the two groups	166
Table 19. Table presenting the comparisons of scores between groups at baseline, 3, 6, and 12/12 timepoints.	168
Table 20. Table to show results from repeated measures tests for all outcomes.	169
Table 21. Table to show average knee joint range of motion in degrees for operated leg at baseline, 3 and 12 months timepoints	175
Table 22. Summary of the content and time (in minutes) of the home visit treatments provided in the trial intervention.	177
Table 23. Summary of the concerns raised by participants and recorded by physiotherapists during the home visit treatments provided in the trial intervention.	178
Table 24. Table to show the total number of non trial physiotherapy visits occurring between discharge from hospital after their knee replacement and one year follow up.	179
Table 25. Table to show the numbers of correct and incorrect guesses of the outcome assessor's beliefs regarding treatment allocation at three month follow up assessments.	224

## LIST OF FIGURES.

	<b>Page Number</b>
Figure 1. Pathogenesis of osteoarthritis with putative risk factors .....	30
Figure 2. Summary of the Sources of Bias when locating and selecting studies for inclusion into a Systematic Review (Summarised using Egger and Davey Smith, 1998).....	48
Figure 3. The Eight Steps for the conduction of a systematic review (Adapted from Egger and Davey Smith, 2001b, p. 25).....	52
Figure 4. Flow diagram to summarise the selection of trials included in the reviews. This diagram is based on the Quorum statement by Moher et al., (1999).....	65
Figure 5. Forest Plot showing Standardised Effect Sizes with Confidence Intervals for .....	74
Figure 6. Forest Plot showing Standardised Effect Sizes with Confidence Intervals for .....	75
Figure 7. Forest Plot showing Weighted Mean Differences with Confidence Intervals for.....	77
Figure 8. Forest Plot showing Weighted Mean Difference (3-4 months) and Standardised .....	79
Figure 9. A summary of imitations of randomised studies (summarised from Black, 1996; Abel and Koch, 1999; McPherson, 1994; Peto and Baigent, 1998; Koes and Hoving, 1998).....	96
Figure 10. An Adaptation of the rehabilitation model suggested by Wade (2005). <i>Italics are used to highlight information regarding the proposed knee replacement trial.</i> .....	107
Figure 11. Summary of MRC Framework for Trials of Complex Interventions (Adapted from the MRC, 2000). .....	110
Figure 12. A Summary Of Strategies Used To Achieve A Theoretical Basis For The Development Of The Trial Intervention (based upon the MRC Framework (2002) ICF (2002) and Wade’s model of rehabilitation (2005). .....	113
Figure 13. CONSORT flow diagram.....	141
Figure 14. Histogram to show the distribution of scores at baseline for the primary outcome, the Oxford Knee Score (n=105). .....	155
Figure 15. Box-whisker plot of the Oxford Knee Score data showing the 2 ½ , 25, 50, 75 and 95 ½ cumulative relative frequencies (centiles) plus one outlier ID 1072. ....	156
Figure 16. SPSS graph to show the one-way repeated measures analysis of variance for the Oxford Knee Score across all timepoints for both groups.....	170
Figure 17. SPSS graph to show the one-way repeated measures analysis of variance for the KOOS Symptoms subscale across all timepoints for both groups. ....	170
Figure 18. SPSS graph to show the one-way repeated measures analysis of variance for the KOOS Pain subscale across all timepoints for both groups.....	171
Figure 19. SPSS graph to show the one-way repeated measures analysis of variance for the KOOS ADL subscale across all timepoints for both groups. ....	171
Figure 20. SPSS graph to show the one-way repeated measures analysis of variance for the KOOS Sports and Recreation subscale across all timepoints for both groups. ....	172
Figure 21. SPSS graph to show the one-way repeated measures analysis of variance for the KOOS Quality of Life subscale across all timepoints for both groups.....	172
Figure 22. Graph to show the one-way repeated measures analysis of variance for the Leg Extensor Press (operated leg) subscale across baseline, 3 and 12 month timepoints for both groups. ....	173
Figure 23. Graph to show the one-way repeated measures analysis of variance for the Timed Walk test across across baseline, 3 and 12 month timepoints for both groups.....	173
Figure 24. Graph to show the one-way repeated measures analysis of variance for the Timed Walk test across across baseline, 3 and 12 month timepoints for both groups.....	174
Figure 25. Histogram showing the number of outpatient physiotherapy visits for the home visit physiotherapy group.....	180
Figure 26. Histogram showing the number of outpatient physiotherapy visits for the usual physiotherapy group. .....	180
Figure 27. Bar chart to show the outcome assessor’s beliefs regarding treatment allocation at three month follow up assessments.....	223
Figure 28. Bar chart to show the outcome assessor’s beliefs regarding treatment allocation at twelve month follow up assessments. ....	224

## **LIST OF PUBLICATIONS AND ABSTRACTS DERIVED FROM THE RESEARCH.**

### **PUBLICATIONS.**

Minns Lowe, C.J., Barker, K. L., Dewey, M., Sackley, C. M. (2007) Effectiveness of physiotherapy exercise after knee arthroplasty for osteoarthritis: systematic review and meta-analysis of randomised controlled trials. *British Medical Journal*; **335**:812-815.

Minns Lowe, C.J., Barker, K. L., Dewey, M., Sackley, C. M. (2009) Effectiveness of physiotherapy exercise following hip arthroplasty for osteoarthritis: a systematic review of clinical trials. *Musculoskeletal Disorders*; **10**:98.

Minns Lowe, C.J., Wilson, M., Barker, K. L., Sackley, C. M. (2009) The feasibility of achieving successful blind outcome assessment in a pragmatic physiotherapy rehabilitation trial. *Submitted for Publication*.

### **ABSTRACTS.**

Minns Lowe, C. J., Wilson, M., Barker, K.L., Sackley, C. M. (2009) Blind outcome assessment in a pragmatic physiotherapy rehabilitation trial. *Ann Rheum Dis*, **68** (Suppl3):778.

Minns Lowe, C. J., Barker, K.L., Holder, R. L., Sackley, C. M. (2009) Effectiveness of physiotherapy following discharge from hospital after primary total knee arthroplasty for osteoarthritis: a randomised clinical trial. *Ann Rheum Dis*, **68** (Suppl3):779.

Minns Lowe, C., Barker, K., Dewey. M., Sackley, C. (2008) A systematic review of physiotherapy exercise programmes following discharge from hospital after total knee arthroplasty. *Australian Journal of Physiotherapy*, **54**:73.

Minns Lowe, C., Barker, K., Dewey. M., Sackley, C. (2008) A systematic review of physiotherapy exercise programmes following discharge from hospital after total knee arthroplasty. Evidence-Based Orthopaedics, *The Journal of Bone & Joint Surgery*, **90-A**: 1172.

Minns Lowe, C., Barker, K., Dewey. M., Sackley, C. (2007) A systematic review of physiotherapy exercise programmes following discharge from hospital after total knee arthroplasty *Clinical Rehabilitation*, **22**: 669-670.

Minns Lowe, C., Barker, K., Sackley, C. (2007) A systematic review of physiotherapy exercise programmes following discharge from hospital after total knee arthroplasty. *Physiotherapy*, **93** (S1): S68.

Minns Lowe, C.J, Barker, K.L., Dewey, M., Sackley, C.M. (2007) A systematic review of physiotherapy exercise programmes following discharge from hospital after total knee arthroplasty. *Annals of the Rheumatic Diseases*, **66** (Suppl II): 117-118 & 640.

Minns Lowe, C. J., Sackley, C. M., Barker, K. (2005) Elective joint surgery in a diagnostic and treatment centre: is a post discharge community physiotherapy intervention effective in improving function and mobility? *Annals of the Rheumatic Diseases*, **64** (Suppl III): 598.

# **1 CHAPTER ONE. INTRODUCTION.**

## **1.1 INTRODUCTION TO THESIS.**

Physiotherapy has traditionally been part of the routine aftercare provided to patients following elective, primary total knee arthroplasty for osteoarthritis. Total knee arthroplasty itself is an increasingly common procedure and the numbers of patients undergoing the procedure is expected to continue to rise. Several issues have recently arisen relating to the rehabilitation of patients following knee joint arthroplasty. The length of hospital stay following surgery has reduced, thus compressing the time available for in-patient rehabilitation. Furthermore, research now indicates that outcome following total knee arthroplasty is less satisfactory than previously assumed, with patients experiencing considerable functional impairment post operatively when compared with their peers.

The effectiveness of physiotherapy following discharge from hospital lacks evaluation. The uncertainty regarding effectiveness subsequently makes it difficult for commissioning organisations, health care practitioners and patients to make decisions regarding rehabilitation and service provision in the UK varies widely. The purpose of this thesis was threefold. Firstly, to evaluate existing evidence regarding post discharge physiotherapy exercise. Secondly, to contribute to the evidence base for this under-researched, but increasingly common, area of physiotherapy practice by developing and evaluating a new physiotherapy intervention. Thirdly to assess the feasibility of achieving blind outcome assessment in a pragmatic rehabilitation trial.

## **1.2 DEVELOPMENTS ALTERING THE THESIS.**

Originally this thesis intended to include both knee and hip arthroplasty patients. The clinical site at which this research was carried out was, at that time, a new Diagnostic and Treatment Centre for hip and knee arthroplasty and the research initially planned to assist the evaluation of this new care approach. Systematic reviews were therefore simultaneously undertaken to evaluate the effectiveness of physiotherapy practice for both knee and hip arthroplasty patients. During this time two developments altered this initial plan. Firstly, the political and health care climate changed and evaluating Diagnostic and Treatment Centres was no longer a priority. Secondly, it became apparent that, whilst there was little research available for the knee, research regarding the effectiveness of physiotherapy following hip arthroplasty was shockingly absent. It was possible to perform a systematic review with meta-analyses for knee joint patients but not for hip patients, where only a narrative review could be performed. This meant that whilst it was possible to use evidence to design a potentially feasible clinical physiotherapy intervention for use in a prospective trial for knee patients, the same could not be said for hip patients for whom research needs to begin at an earlier level. The decision was made therefore to tighten the focus of this Ph.D. thesis to knee arthroplasty patients only. This decision was also made to improve the flow of the thesis which thus concentrated on knee joint arthroplasty rehabilitation. The hip systematic review was completed (see list of publications) but is not included in this thesis.



### **1.3 AIMS AND HYPOTHESES OF THIS THESIS.**

#### **1.3.1 PART ONE. THE SYSTEMATIC REVIEW.**

##### **1.3.2 Background.**

Initial literature reviews indicated that much of the published literature surrounding physiotherapy following joint arthroplasty consisted of descriptions of hospital care pathways and non evidence based opinions regarding rehabilitation. However, a change could be seen in recent years with the publication of clinical trials exploring physiotherapy interventions after discharge from hospital. No systematic review of these trials had yet been performed and it seemed appropriate for the first part of this thesis to consist of an evaluation of these trials.

##### **1.3.2.1 Aim of the Systematic Review.**

A systematic review of the literature evaluating the effectiveness of post operative physiotherapy exercise following elective joint arthroplasty was undertaken to determine the extent of previous research, to check the value of the future research in to the topic area, and to inform the development of a future trial intervention and the choice of appropriate outcomes for such a trial.

##### **1.3.2.2 Objective for the Systematic Review.**

The objective for the systematic review is as follows: to what extent is post discharge physiotherapy exercise effective, in terms of improving functional activities of daily living,

walking, range of joint motion, muscle strength and quality of life, for osteoarthritic patients following primary unilateral elective total knee arthroplasty?

### **1.3.3 PART TWO. THE RANDOMISED CLINICAL TRIAL.**

#### **1.3.3.1 Background.**

The systematic review identified that functional interventions seem more promising than traditional programmes in improving short term function, range of joint motion and quality of life. The review highlighted the need for further trials evaluating physiotherapy following knee replacement, a need identified in both a nationwide survey of physiotherapy rehabilitation following knee replacement in Australia (Naylor *et al.*, 2006) and a subsequent review (Dauty *et al.*, 2007). It was decided therefore to undertake a phase II randomised clinical trial to evaluate whether such a functional exercise intervention provided to patients following discharge after primary unilateral total knee arthroplasty was of benefit in terms of improving terms of improving both self report and objective functional outcomes.

#### **1.3.3.2 Aim of Trial.**

The aim of the trial was to compare the effectiveness of a post discharge physiotherapy intervention in improving patient self report and objective functional outcomes versus usual care for osteoarthritis patients undergoing primary total knee arthroplasty.

### **1.3.3.3 Hypotheses for the Trial.**

The hypotheses for the trial are now presented.

*Alternate Hypothesis for the Primary Outcome:* Participants receiving two additional physiotherapy home visits after discharge from hospital following knee joint arthroplasty will show greater improvement in Oxford Knee Scores than participants receiving usual physiotherapy care.

*Null Hypothesis for the Primary Outcome:* There will be no statistically significant difference in Oxford Knee Outcome Scores at 3,6, and 12 months post operation time points between participants receiving two physiotherapy home visits and participants receiving usual physiotherapy care.

*Alternate Hypotheses for the Secondary Outcomes:* For each of the secondary outcomes, namely, the Knee injury and Osteoarthritis Outcome Score (KOOS), the Leg Extensor Press, the Timed Walk test and the Timed Sit to Stand test the follow alternate hypothesis will be tested.

Participants receiving two additional physiotherapy home visits after discharge from hospital following knee joint arthroplasty will show greater improvement in the secondary outcome scores than participants receiving usual physiotherapy care.

*Null Hypotheses for the Secondary Outcomes:* For each of the secondary outcomes, namely, the KOOS, the Leg Extensor Press, the Timed Walk test and the Timed Sit to Stand test the follow null hypothesis will be tested.

There will be no statistically significant difference in the secondary outcome scores at 3,6, and 12 months post operation time points between participants receiving two physiotherapy home visits and participants receiving usual physiotherapy care.

### **1.3.4 PART THREE. THE EVALUATION OF BLIND OUTCOME ASSESSMENT.**

#### **1.3.4.1 Background.**

During the systematic review it became apparent that many systems / scores used to assess the quality of trial reports immediately denigrate and downgrade a trial's quality if it is not a double blinded trial. This occurs even when double blinding is inappropriate or impossible to achieve within a trial, as in the trials included in the systematic trial in Chapter 3, and even though the vast majority of reports provide no information whatsoever regarding the success or otherwise of their blinding procedures. It cannot be assumed that the use of procedures to promote blinding in trials lead automatically to a successfully blinded trial, even what is meant by a successfully blinded trial appears unknown, and yet the presence of blinding procedures are still being used as a key indicator of trial quality.

For pragmatic trials, where double blinding is often inappropriate, the emphasis has been upon ensuring blind outcome assessment. The systematic review shows blind outcome assessment procedures were used in the component trials although, again, no details about the feasibility and success of blind outcome assessment were included in any of the trial reports. Since no research could be found regarding the feasibility of achieving successful blind outcome assessment in non pharmacological trials, or what constitutes success, it was considered appropriate to explore this topic further during this PhD by measuring to what

extent blind outcome assessment occurred in the trial mentioned above and by exploring the circumstances surrounding incidents of unblinding.

#### **1.3.4.2 Aim of Study.**

The aim of this study was to explore the feasibility of achieving blind outcome assessment in a pragmatic physiotherapy rehabilitation trial involving older people and to contribute to the limited available knowledge in this area.

#### **1.3.4.3 Objectives of the Study.**

The objectives of the study were as follows:

1. To record and present the number of instances of unblinding occurring during the trial.
2. To fully document details surrounding each instance of unblinding and to present a content analysis of the results.

### **1.4 OVERVIEW OF CHAPTERS.**

Chapter two provides a summary of the background literature regarding osteoarthritis of the knee, the treatment and management of osteoarthritis of the knee, the prevalence of total knee arthroplasty and rehabilitation and outcome following arthroplasty.

Chapter three presents the systematic review used to evaluate the effectiveness of post discharge physiotherapy exercise on function, walking, range of motion, quality of life and muscle strength, for osteoarthritic patients following elective primary total knee arthroplasty. It includes a background literature review and a justification for the choice of methodology,

the review protocol, the results of the review and the discussion sections. The discussion section includes a discussion of the implications for clinical practice and identifies areas of future research.

Chapter four presents the randomised clinical trial to compare the effectiveness of an innovative post discharge physiotherapy intervention in improving patient function versus usual physiotherapy for osteoarthritis patients undergoing primary total knee arthroplasty. The chapter includes a background literature review, a justification for the choice of methodology, a section describing the development and rationale of the trial intervention, plus the trial protocol and results sections. Following this the discussion section is presented, this includes implications for clinical practice and identifies areas of future research.

Chapter five presents a study to explore the feasibility of achieving blind outcome assessment in a pragmatic physiotherapy rehabilitation trial involving older people. It includes a background literature review and a justification for the choice of methodology, plus the study protocol and results. The discussion section follows next and again this includes a discussion of the implications for clinical practice and identifies areas of future research.

Chapter six summarises the main findings and implications, the future research suggestions and the conclusions of this thesis.

## **2 CHAPTER TWO. BACKGROUND LITERATURE REVIEW.**

### **2.1 CHAPTER OVERVIEW.**

This chapter commences with definitions of osteoarthritis and will then provide current opinions regarding its aetiology. The prevalence of osteoarthritis of the knee and the signs and symptoms and current concepts of treatment of this condition will then be presented. This is followed by a section on knee joint arthroplasty; including factors known to effect outcome and the role of physiotherapy within post operative rehabilitation.

### **2.2 DEFINITIONS OF OSTEOARTHRITIS.**

There are multiple definitions of osteoarthritis: a reflection of both the lack of consensus and the development of understanding regarding the condition. The condition is one of great antiquity, with examples of the condition being found in fossils dating back 100 million years, with the term itself dating from 1890 (Dequeker and Luyten, 2008). The term osteoarthritis remains problematic; some believe it to be a misnomer since it implies an inherently inflammatory condition and suggest preferred alternative terms such as osteoarthrosis and degenerative joint disease (Dequeker and Luyten, 2008).

Definitions have ranged from the rather vague:

Clinical and pathological outcome of a range of disorders that results in structural and functional failure of synovial joints

Hunter and Felson, 2006

which has the benefit of including function, to the more precise:

a non-inflammatory disorder of movable joints characterised by deterioration and abrasion of articular cartilage, and also by formation of new bone at the articular surface and subchondral bone involvement.

Dequeker and Luyten, 2008

One of the most recent inclusive definitions was developed at a workshop of experts (Kuettner and Goldberg, 1995 cited in Brandt *et al.*, 2008) and is as follows:

Osteoarthritis is a group of overlapping distinct diseases which may have different etiologies, but with similar biologic, morphologic, and clinical outcomes. The disease processes not only affect the articular cartilage, but involve the entire joint, including the subchondral bone, ligaments, capsule, synovial membrane, and periarticular muscles. Ultimately, the articular cartilage degenerates with fibrillation, fissures, ulceration, and full thickness loss of the joint surface. OA diseases are a result of mechanical and biologic events that destabilise the normal coupling of degradation and synthesis of articular cartilage of chondrocytes and extracellular matrix, and subchondral bone. Although they may be initiated by multiple factors, including genetic, developmental, metabolic, and traumatic, OA tissues involve all of the tissues of the diarthrodial joint. Ultimately, OA diseases are manifested by morphologic, biochemical, molecular, and biomechanical changes of both cells and matrix which lead to a softening, fibrillation, ulceration, loss of articular cartilage, sclerosis and eburnation of subchondral bone, osteophytes, and subchondral cysts. When clinically evident, OA diseases are characterised by joint pain, tenderness, limitation of movement, crepitus, occasional effusion, and variable degrees of inflammation without systemic effects.

Even such comprehensive definitions are incomplete; not mentioning important factors such as joint biomechanics or emphasizing that osteoarthritis reflects the failed repair of damage caused by excessive mechanical stress on joint tissues (Brandt *et al.*, 2008). Although the term osteoarthritis also encompasses patients with very rare systemic diseases (Brandt *et al.*, 2008) such subjects are not the subject of this thesis and will not be further discussed.



For the purposes of this thesis symptomatic osteoarthritis is defined as a painful disease whereby processes of aberrant repair, as discussed above, involve the entire joint and affect the whole person. Sufferers experience pain, impairments, limitations in functional abilities and activities plus restricted participation in activities and behaviours of their choice. The disease not only affects physical health but can, and often does, impact upon emotional health and prevent sufferers from living the type of lives they wish to lead within their society.

### **2.3 SIGNS AND SYMPTOMS OF OSTEOARTHRITIS OF THE KNEE.**

Osteoarthritis of the knee may cause pain and functional disability (Jordan *et al.*, 2003). There is a unexplained lack of correlation between disease severity and the level of reported disability in pain and the source and causes of pain are complex and not well understood (Hunter *et al.*, 2008). Whilst pain is not generated by the aneural cartilage there are many other structures that may contribute to pain, for example subchondral bone, ligaments, the periosteum, synovium and joint capsule are all richly innervated (Dieppe and Lohmander, 2005). There is also no simple definitive test for osteoarthritis; severe osteoarthritis may be detected by radiography but radiographic findings relate poorly to self reported pain, function, muscle power and disability (Rogers *et al.*, 2004; Barker *et al.*, 2004; Creamer *et al.*, 2000; Hannan *et al.*, 2000).

Hunter *et al.*, (2008) reviewed the main symptoms of osteoarthritis as follows:

1. Activity related or mechanical pain usually of insidious onset (which may occur at rest in severe disease); often non localised deep aching pain and associated with increased age. Biologic, psychologic and social factors all seem involved in pain creation and perception.

There is also some evidence that women report greater pain levels than men (Affleck *et al.*, 1999).

2. Limited function.
3. Stiffness; usually shortlived after inactivity.
4. Joint instability.

Hunter *et al.*, list additional possible symptoms as reduced range of joint motion, joint deformity, swelling not associated any systemic causes, pain related psychological distress, altered gait and muscle atrophy or weakness.

Gooberman-Hill *et al.*, (2007) explored pain experiences with people (n=28) suffering from clinical knee and hip osteoarthritis. They found that pain was generally experienced as intermittent (or in bouts) and was variable by activity, time of day, day by day and month by month. One participant, however, described constant pain. Pain was also often present elsewhere in the body and was inextricably linked with function, leading to avoidance or adaptation of activities.

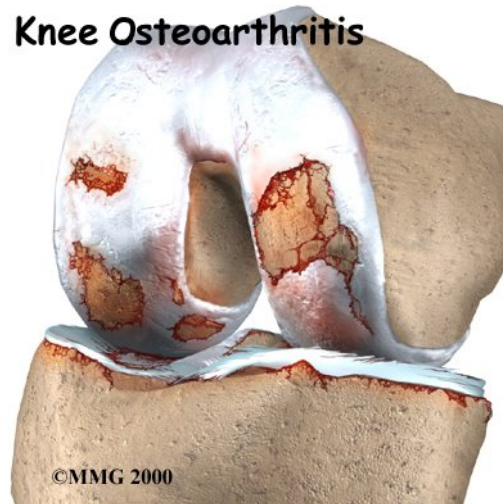
Symptomatic osteoarthritis of the knee can lead to significant disability. Creamer *et al.*, (2000) developed a model of determinants of disability in symptomatic knee osteoarthritis. They used demographic data and self report function and pain scores and found that pain severity, obesity and helplessness accounted for 59.9% of variance in the Western Ontario and McMaster Universities OA Index disability scores. Further evidence for significant disability from osteoarthritis is provided by Fautrel *et al.*, (2005). Their French nationwide survey of over 10,000 patients with osteoarthritis found that limitations in performing activities of daily living, such as mobility outside the home, house chores and work duties, affect over 80% of patients . For the 3,247 patients with knee osteoarthritis, 72.8% of patients reported mobility problems when outside and, of the minority still working,

65.7% reported their osteoarthritis created occupational limitations. The substantial functional impact of osteoarthritis therefore seems evident.

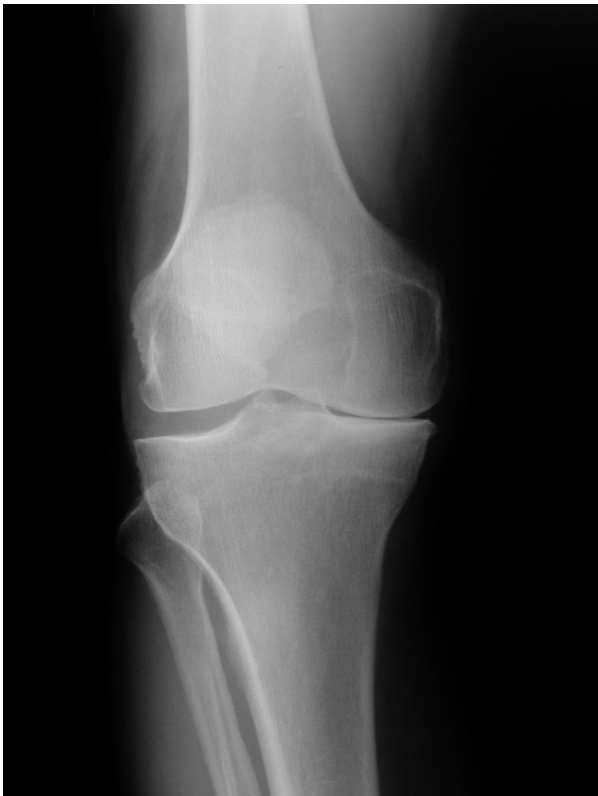
## **2.4 AETIOLOGY OF OSTEOARTHRITIS OF THE KNEE.**

Osteoarthritis is a multifactorial process caused by “aberrant local mechanical factors acting within the context of systemic susceptibility” (Hunter and Felson, 2006). It is a complex disease, since its aetiology bridges both biochemistry and biomechanics, and it remains unclear whether osteoarthritis is a single disease or many diseases which subsequently demonstrate a similar final pathway (Felson *et al.*, 2000a). Whilst traditionally osteoarthritis was considered a disease of articular cartilage with hyaline articular cartilage loss (Felson & Neogi, 2004), the contemporary view is that osteoarthritis involves the entire joint organ of subchondral bone, synovium, menisci, ligaments, periarticular muscle and capsule and that the disorder is characterised by tissue response in addition to tissue destruction (Hunter & Felson, 2006; Brandt *et al.*, 2006; Dequeker and Luyten, 2008). Abnormal joint mechanics “provoke biological effects that are mediated biochemically – for example through cytokines, matrix-degrading enzymes and toxic oxygen radicals” (Brandt *et al.*, 2006) causing a complex aetiology. In the knee the disease is primarily characterised by focal loss of articular cartilage plus marginal and central new bone formation (Jordan *et al.*, 2003). The loss of cartilage is clearly shown in the illustration below. The subsequent X ray pictures of knee osteoarthritis also illustrate these disease processes.

## Knee Osteoarthritis



**Illustration 1. Drawing of an osteoarthritic knee (eorthopod.com/images).**  
Showing joint space narrowing, eroding meniscii and cartilage, and areas of exposed bone.



**Illustration 2. Radiograph showing an anterior-posterior view of an osteoarthritic knee joint with medial joint space narrowing and subchondral bone thickening.**



**Illustration 3. Radiograph showing a lateral view of an osteoarthritic knee joint.**

While severe joint injury may be sufficient to cause osteoarthritis, Dequeker and Luyten (2008) comment that many of the radiographic and clinical features of the disorder are due to attempted repair processes occurring within the joint, the disease is often produced after an interplay of both local and systematic factors (Felson *et al.*, 2000). Brandt *et al.*, (2008) concur that osteoarthritis reflects processes occurring in joints attempting to contain damage.

In a comprehensive editorial reviewing recent evidence regarding the aetiology of osteoarthritis Brandt *et al.*, (2006) usefully summarise possible effects of osteoarthritis on the entire joint organ as follows:

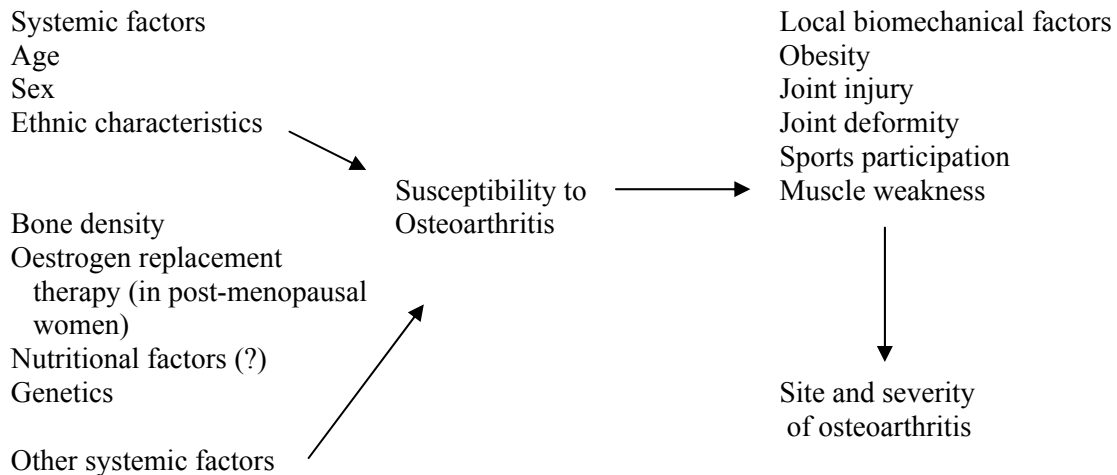
- a) Excessive loading and abnormal mechanical stresses can damage hyaline articular cartilage. Alterations in gene expression can be caused leading to the degradation of cartilage. In addition when other joint tissue damage also occurs, such as joint capsule/synovium damage, higher levels of cartilage degrading chemicals may be produced than with cartilage damage alone.
- b) Ligament damage resulting in ligamentous laxity may lead to joint instability and osteoarthritis.
- c) Periarticular muscle weakness and non optimal loading on joint cartilage, for example through impulsive or unexpected loads, may also lead to osteoarthritis.
- d) Reduced proprioceptive ability may predispose or be linked to the development of osteoarthritis.
- e) Subchondral bone is thickened. This is thought to be due to altered biomechanics either leading to increased bone turnover and the reactivation of the secondary centre of ossification or to an increase in the stiffness of subchondral bone which makes the bone less capable of attenuating and distributing the load throughout the joint. This reduced capability increases stresses in the overlying articular hyaline cartilage which subsequently deteriorates.
- f) Meniscus subluxation may account for a large amount of the reduction of knee joint space rather than hyaline cartilage thinning.

In addition, Felson and Neogi (2004) describe the formation of osteophytes as a further characteristic of osteoarthritis. These are outgrowths of cartilage at the joint margins, which secondarily ossify. Sclerosis then probably occurs because the localised mechanical stress placed on the bone leads to bony proliferation. Brandt *et al.*, (2008) disagree, using a slide of a Heberden's node to show that osteophytes are not necessarily pathognomonic for osteoarthritis.

Mansell *et al.*, (2007) discuss the hypothesis that it is the increasingly thickened and stiffened subchondral bone that leads to articular cartilage damage during repetitive loading, believing that, whilst questions remain regarding the stimulation of such changes in bone tissue, osteoarthritis involves a metabolic dysfunction of bone rather than “wear and tear” of cartilage. An osteological study by Rogers *et al.*, (2004) supports the theory that osteoarthritis is part of a systemic disorder of bone. This study of 563 skeletons found that evidence of osteoarthritis in classic sites, such as the knee and hip, is associated with widespread skeletal osteoarthritis type changes. This challenges the division of osteoarthritis patients into monoarticular and generalised osteoarthritis groups, indicating that osteoarthritic joint changes are more likely to be widespread with some joints remaining asymptomatic. Similarly this also challenges the previous distinction between primary or secondary osteoarthritis, or the presence of absence of any obvious cause for the condition which the multifactorial nature of the disease makes complex (Dieppe and Lohmander, 2005).

Systemic factors which appear to increase the vulnerability of joints to symptomatic osteoarthritis include increasing age, female gender, possible nutritional deficiencies and a major (as yet unidentified) genetic component whilst local factors include misalignment, muscle weakness and damage/alterations to the structural integrity of the joint organ (Hunter and Felson, 2006). Occupational factors, such as repetitious kneeling, squatting and heavy lifting are also associated with increased susceptibility to knee osteoarthritis (Felson *et al.*, 2000a). Furthermore injury and/or obesity can effect joint loading and adversely influence the development and progression of the condition (Hunter and Felson, 2006; Messier, 2008). A summary of risk factors is shown in the figure below. Since certain factors, such as obesity, are more clearly associated with knee osteoarthritis than hip osteoarthritis, some feel this adds support to the argument that osteoarthritis is likely to be several distinct disease entities

(Felson *et al.*, 2000a). Given Rogers *et al.*, (2004) study this might perhaps be linked to how/why joints become symptomatic rather than non generalised osteoarthritis *per se*.



**Figure 1. Pathogenesis of osteoarthritis with putative risk factors**

This figure is Figure 2 from Felson *et al.*, 2000a [a modified version of the original figure by Dieppe, P. The classification and diagnosis of osteoarthritis. In: Kuttner, K., Goldberg, V. eds. Osteoarthritis Disorders. Rosemount, IL: American Academy of Orthopaedic Surgeons 1997]. Permission to reproduce this figure has been obtained from the Annals of Internal Medicine (see Appendix 1).

## **2.5 THE PREVALENCE OF OSTEOARTHRITIS OF THE KNEE AND THE RATE OF PRIMARY TOTAL KNEE JOINT REPLACEMENT IN THE UK.**

It is difficult to measure the prevalence of knee osteoarthritis since disparities exist between patient reported symptoms, objective clinical diagnosis and different systems of radiographic evidence (D’Ambrosia, 2005). A recent estimate is that osteoarthritis of the knee results in



disabling knee symptoms in an estimated ten per cent of people aged over fifty five, a quarter of whom are severely disabled (Peat *et al.*, 2001). In an older hallmark study, the Framingham Osteoarthritis study, 1,424 knee radiographs were graded for osteoarthritis changes, which provided prevalence estimates ranging from 11.5% in participants aged under 70 years to 19.4% in those aged over 80 (Felson *et al.*, 1987). The different prevalences obtained via radiography may be partially due to some people having radiographic changes whilst remaining asymptomatic (Hannan *et al.*, 2000). Since osteoarthritis increases with age it is recognised that the incidence will increase in the future in line with the aging population (Felson *et al.*, 2000a).

The number of primary total knee joint replacement procedures is increasing, which may be both in response to the aging population and to the previous underprovision of knee arthroplasty (Jüni *et al.*, 2003). The rate of primary total knee joint replacement procedures in England rose by up to 63% between 1991 and 2000 (Dixon *et al.*, 2004) and has tripled in Scotland in the last 15 years (Scottish Arthroplasty Project Annual Report 2008). In 2007, 55,091 primary total knee replacements were undertaken in England and Wales for osteoarthritis; 97% of all primary knee replacement procedures (National Joint Registry 5<sup>th</sup> Annual Report). Knee replacements were more common for women (58%) and for older people (average age 70.5). Knee replacement surgery is also more common for people not in the most deprived fifth of the population as measured via the Townsend index of population (Dixon *et al.*, 2004). In Scotland in 2006/7 the number of total knee replacements overtook the number of hip replacements for the first time with 5799 total knee procedures being undertaken for osteoarthritis (Scottish Arthroplasty Project Annual Report 2008). Latest figures for Northern Ireland from the of Department of Health, Social Services and Public

Safety, obtained using the OPCS codes W40-42 for knee replacement, provisionally suggest that 1429 knee replacements were carried out in 2007/8 (Stewart, 2009).

## **2.6 CURRENT CONCEPTS FOR TREATMENT OF OSTEOARTHRITIS OF THE KNEE.**

A recent fundamental shift has been from viewing osteoarthritis as a passive and degenerative disorder, for which little can be done, to viewing it as an active disease process susceptible to modification by mechanical and biochemical treatments (Dequeker and Luytens, 2008). The importance of restoring abnormal joint mechanics to a physiologic range to enable healing to occur has been emphasised by Brandt *et al.*, (2008).

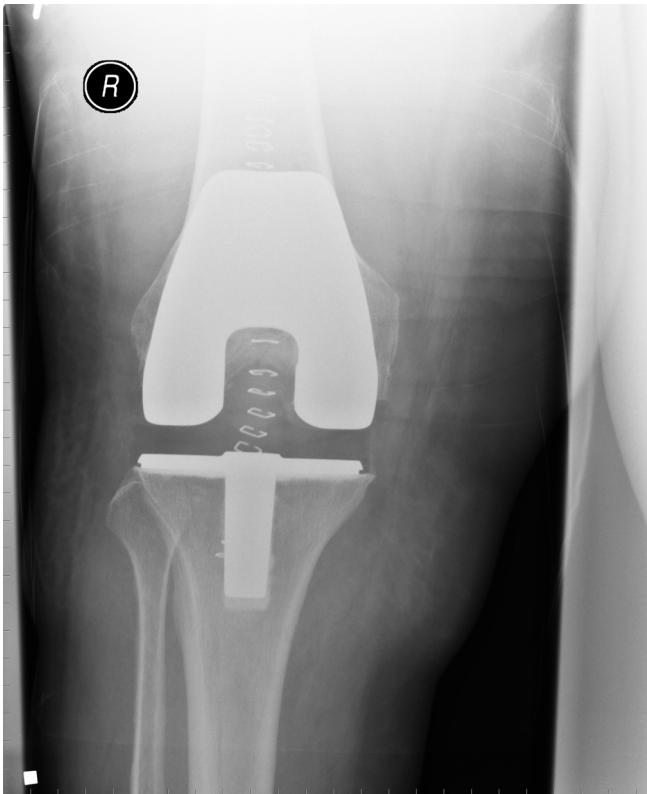
Even recent commentaries upon osteoarthritis however present very different treatments aims. Some authors advocate the use of therapeutic strategies primarily aimed at reducing pain and improving function (Hunter and Felson, 2006) while others advocate the mechanical or biochemical manipulation of the disease processes itself (Dequeker and Luyten, 2008). The emphasis upon treating cartilage alone by developing / seeking chondroprotective drugs has been recently criticised with a plea to direct attention to the entire joint organ and joint biomechanics (Brandt *et al.*, 2006 and 2008). Similarly, the EULAR evidence based recommendations cover both non pharmacological treatments, such as exercise and lifestyle changes, pharmacological treatments, such as paracetamol and NSAIDS, and invasive interventions such as surgery (Jordan *et al.*, 2003). These guidelines state that, although the evidence available to support treatment efficacy remains variable, the optimal management of osteoarthritis of the knee includes both non pharmacological and pharmacological treatments. The aims of disease management listed are patient education, pain alleviation, the reduction of

disability and improvement of function and the prevention/retardation of the disease and management should be tailored according to a patient's risk factors. Further evidence that exercise therapy is considered effective for patients with knee osteoarthritis has been provided by a review of three systematic reviews (Smidt *et al.*, 2005). The muscles around the knee produce movement, absorb limb loading forces and provide dynamic joint stability around the knee: muscle weakness is both a risk factor for, and a potential result of, osteoarthritis of the knee (Bennell *et al.*, 2008). This review of the role of muscle in the genesis and management of knee osteoarthritis usefully highlights the strong evidence that muscle strength can be improved by muscle strengthening programmes in people with knee osteoarthritis. Such evidence has led to the call for individualised holistic treatment plans with the recommendation that exercise should be a core treatment for people with osteoarthritis (National Institute for Health and Clinical Excellence, 2008).

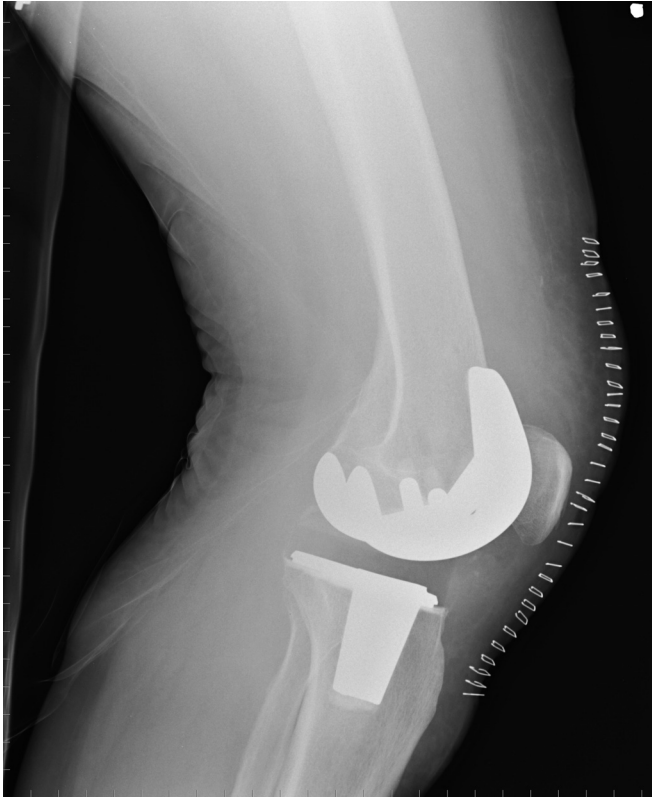
## **2.7 JOINT ARTHROPLASTY.**

Joint arthroplasty is an intervention for those patients who have severe disease, i.e. daily severe pain plus radiographic evidence, who have not responded satisfactorily to other treatment modalities (Jordan *et al.*, 2003). Occasionally progressive deformity or instability may be the prime indicator (British Orthopaedic Association and the British Association for Surgery of the Knee, 1999). Patient motivation, age, health seeking behaviours, general practitioner referral pattern, the presence of co-morbidities, reluctance/willingness to undergo surgery, functional impairment and socio-economic status are other factors influencing whether or not people undergo arthroplasty surgery (Dieppe *et al.*, 1999; Mancuso *et al.*, 1996; Juni *et al.*, 2003).

The procedure involves an anterior incision to expose the distal end of the femur, the proximal end of the tibia and the posterior articular surface of the patella; followed by the insertion of the prostheses (with/without cement), any necessary balancing of the medial and lateral soft tissue structures and correction of deformities, then ensuring satisfactory patella tracking and the integrity of the extensor mechanism of the knee prior to wound closure (British Orthopaedic Association and the British Association for Surgery of the Knee, 1999). The following radiographs show a knee arthroplasty *in situ*.



**Illustration 4. Radiograph showing an anterior posterior view of a total knee joint arthroplasty.**



**Illustration 5. Radiograph showing lateral view of a total knee arthroplasty.**

As a guide to how well knee replacements perform over time, one of the largest and most recent survivorship analyses of 11,606 primary total knee arthroplasties showed 92% of knees surviving to 10 years, falling to 78% at twenty years (Rand *et al.*, 2003). Patients under 55, considered more active than older patients, fared worse at 10 years (83%) than those aged over 70 (94%). Men fared worse (88%) at 10 years than women (93%). Type of prosthesis, use of cement, and diagnosis also influenced survivorship. Although long term figures are not yet available from the National Joint Registry the short term trends have been for a higher revision rate in those aged under 65, with a non statistically significant higher rate of revision amongst women than men (National Joint Registry 4<sup>th</sup> Annual Report p. 89-93).

It has been pointed out that, whilst an established treatment, there are few trials of knee joint arthroplasty (Dieppe *et al.*, 1999) and no trials yet comparing joint arthroplasty

with non-surgical interventions (Jordan *et al.*, 2003). There has also been wide variation in the types of prosthesis used (Dieppe *et al.*, 1999). It has also been noted that for the younger population, for whom prosthesis durability is an issue, alternative and less invasive treatments to total arthroplasty (such as osteotomy or unicompartmental procedures) appear to have a role requiring further research (Richmond, 2008).

## **2.8 OUTCOME AFTER JOINT ARTHROPLASTY.**

Total joint arthroplasties have long been considered effective interventions to improve pain and function in patients, with few serious adverse outcomes and low mortality rates (Jones *et al.*, 2007). Specific knee arthroplasty postoperative complications known to hinder recovery, and their rates, will be presented in Chapter four. A relatively early systematic review reported good or excellent outcome for pain and mobility in 89.3% of patients (Callahan *et al.*, 1994). However, the choice of outcome measure influences the results obtained; only revision and prosthesis failure rates were previously reported but, as survivorship rates became so successful, patient-centred outcomes began to be used more frequently (Woolhead *et al.*, 2005; Wylde *et al.*, 2007). The latter reveals that outcome is multifaceted. Jones *et al.*, (2007) state that 15-30% of patients report little/no improvement, or dissatisfaction, post surgery. Problems include persistent pain, impairment and functional limitations (Westby, 2008): the determinants of poor pain and functional outcomes are still largely unknown and it has been suggested that psychological factors may be influencing outcome in addition to physical, socio-demographic, medical, surgical and prosthetic-related factors (Jones *et al.*, 2007; Wylde *et al.*, 2007).

Recent qualitative research by Woolhead *et al.*, (2005) has helpfully further contributed to knowledge regarding outcome by exploring patients' perspectives of outcome

following knee arthroplasty. Most patients in the study provided contradictory accounts: whilst reporting a good outcome from surgery they went on to recount continuing pain and mobility problems. Patients struggled to make sense of these continuing difficulties and tried to take personal responsibility for them. This keenness to take responsibility for outcome was also noticed in an earlier qualitative study with patients undergoing orthopaedic surgery (Edwards, 2002). A further qualitative study after joint arthroplasty found that incongruence between patient expectations and reality regarding recovery was a source of distress (Showalter *et al.*, 2000). It should be recognised that the outcomes used in many studies will not capture or reflect these complex issues and that patients have different opinions regarding outcome after arthroplasty than health care professionals, particularly when the patient is dissatisfied with their outcome (Wright *et al.*, 1994; Lieberman *et al.*, 1996).

It should also be recognised that patients who have undergone knee arthroplasty experience substantial functional impairment when compared with their age and gender matched peers (Noble *et al.* 2005; Walsh *et al.*, 1998). Noble *et al.*, found that the gap between post arthroplasty and control groups widened as activities became more demanding and that patients shifted their attention to less demanding pursuits upon discovering the limitations their joint arthroplasty imposes upon them. The authors conclude that patients should be warned that their knee pain is likely to be improved after arthroplasty, as the literature indicates (Wylde *et al.*, 2007), but that their knee will still probably not function as well as it did prior to the onset their arthritis.

Small scale research shows that persistent quadriceps weakness post knee arthroplasty is associated with poorer functional outcome (Silva *et al.*, 2003). Another recent small study (n=56) demonstrated that focused physiotherapy can improve outcome (as measured by the Knee Society score) and patient satisfaction at 24 months post operation amongst patients

reporting functional problems at least two months post knee joint arthroplasty (Ulrich *et al.*, 2007). Quadriceps muscle weakness and flexion contractures were the most common problems observed and treated. Knee muscle strength has previously been shown to be up to 30-40% weaker for knee arthroplasty patients at 1 year post operation when compared to controls with no knee disease (Walsh *et al.*, 1998).

Preoperative mobility appears to affect the recovery of mobility following knee arthroplasty: in a group of 76 patients recovery of walking speed and stair climbing was predicted by preoperative total power and, to a lesser extent, by body mass index (Lamb and Frost, 2003). Interestingly, the patients in the lowest tertile of preoperative leg power experienced the greatest recovery rates but did not achieve the levels of mobility reached by those starting with higher power within the 6 months follow up period. Again, walking speed and stair climbing speeds are known to remain far slower for knee arthroplasty patients at 1 year post operation when compared to controls with no knee disease (Walsh *et al.*, 1998). The existence of other joint disease is also known to significantly and adversely affect post operative recovery mobility and function (Naylor *et al.*, 2008). Mobility is impaired and recovery times are slowed.

Overall therefore, outcome following knee joint arthroplasty appears generally effective in terms of achieving pain relief but less effective in terms of regaining pre-arthritis levels of function.

## **2.9. REHABILITATION.**

This section will outline physiotherapy rehabilitation procedures for knee joint arthroplasty. It will not include an exploration of the effectiveness of physiotherapy exercise programmes



following discharge since this subject is the focus of a systematic review in chapter three. Whilst comprehensive narrative reviews exist for rehabilitation following total hip replacement (for example, Brander and Mullarkey, 2002) the picture for total knee arthroplasty is less clear. There are no evidence based clinical guidelines post knee arthroplasty (Westby *et al.*, 2008). As Naylor *et al.*, (2006) phrase it “the evidence base regarding total knee replacement rehabilitation is somewhat fragmented” with considerable evidence for the use of acute postoperative cryotherapy and continuous passive movement but little for the treatments of choice such as exercise or hydrotherapy or for the progression of functional activities.

Pre-operative home physiotherapy remains rare in the UK with one trial demonstrating its lack of effect upon patient perceived health outcomes (Mitchell *et al.*, 2005). Similarly, a review has found physical training before knee arthroplasty to be of no value (Dauty *et al.*, 2007).

Post-operatively, the main aims of physiotherapy rehabilitation following knee arthroplasty have been summarised as follows: to reduce pain, to maintain/increase joint range of motion, to maintain/improve muscle strength (particularly the quadriceps and gluteal muscles), to maximise functional ability (including improving/maintaining proprioception), to support/advise and educate patients as necessary and to encourage self-care and self management (Coutts, 2005 p 251-262). The prevention of postoperative complications has been added to this list by Jones *et al.*, (2005).

In-patient rehabilitation concentrates on achieving sufficient functional independence for discharge and on improving knee range of motion and muscle strength (Oldmeadow *et al.*, 2002). The need to prepare patients for hospital discharge as soon as possible has been emphasized (Lenssen *et al.*, 2006) and there is pressure on clinicians to decrease and

minimise length of stay (Oldmeadow *et al.*, 2002). Indeed the goal of many in-patient pathways is to reduce costs and length of stay in hospital without compromising upon patient outcome and many pathways are considered successful in achieving this goal (Kim *et al.*, 2003; Husted and Holm, 2006; Iyengar *et al.*, 2007). The average UK length of hospital stay for total joint arthroplasties other than hips in 2005-6 was 8.2 days (Hospital Episode Statistics Online).

Existing rehabilitation pathways and protocols show wide variation (Jones *et al.*, 2005). Generally, early mobilisation has been promoted as the gold standard approach for the achievement of functional mobility (Roos, 2003). Patients usually receive inpatient physiotherapy daily and are also advised regarding additional exercise and mobility activities to continue between sessions; the use of multiple sessions per day may occur but a trial comparing once or twice daily sessions has questioned the benefit of multiple daily sessions (Lenssen *et al.*, 2006). Further adjuncts to exercise and mobility may be used if length of stay, staff time and resources permit. An example would be the use of continuous passive motion machines which, if tolerated by patients and available, can benefit range of joint motion and analgesia (Milne *et al.*, 2003).

The dosage and management of physiotherapy following the acute postoperative phase after surgery does vary widely (Roos, 2003; Jones *et al.*, 2005), with financial reimbursement schemes seeming to influence practice (Lingard *et al.*, 2000; Roos, 2003). The amount of rehabilitation in some countries far exceeds the UK. A recent Netherlands, non assessment blinded trial, compared 3 weeks intensive rehabilitation in a resort (up to 4 hours with a physiotherapist per day) compared with 'usual care' of local physiotherapy or admission into a nursing home for rehabilitation (Bulthuis *et al.*, 2007). The usual care therefore also appears intensive which may explain why, although the resort group improved more quickly,

there was no significant difference at a year. The trial (n=114) may also have been underpowered.

In the absence of evidence, Naylor *et al.*, (2006) undertook a nationwide survey of current physiotherapy practice following knee arthroplasty in Australia. Consistently they found that physiotherapists provided gait retraining and exercise prescription, with 57% of programmes including functional activities and 83% including exercises, and with most programmes operating on a one-to-one basis. Unlike an earlier USA consensus study (Enloe *et al.*, 1996), this Australian survey therefore does now include functional exercises. The mean duration of post discharge programmes offered was 5.6 weeks and programmes were offered by the majority of providers (88%). The authors conclude that the survey suggests that the exercise programmes lack may adequate intensity to lead to optimal recovery, a concern also raised by Westby *et al.*, (2008).

In the UK the nature of post operative rehabilitation practices sounds similar to the Australian survey except that less UK patients are less likely to be referred for post discharge physiotherapy due to its disputed value (Rajan *et al.*, 2005). Anecdotally we know of many other Trusts who do not routinely arrange post operative physiotherapy.

In addition to the above, evidence demonstrates that progressive quadriceps muscle strengthening improves functional outcome following knee arthroplasty (Mizner *et al.*, 2005; Ulrich *et al.*, 2007; Petterson *et al.*, 2009). This is not surprising given that muscles around the knee produce movement, absorb limb loading and provide dynamic joint stability (Bennell *et al.*, 2008). Trials have shown quadriceps strength to be a stronger predictor of timed up and go, stair climbing and six minute walk test performance than pain and range of joint motion (Mizner *et al.*, 2005; Petterson *et al.*, 2009). Participants receiving the muscle strengthening programme were significantly stronger and more able during the functional

tests than those receiving standard care (range of motion and non weightbearing exercises) (Pettersen *et al.*, 2009). It is worth mentioning that the mean number of physiotherapy outpatient visits in this programme was 17 during the six week intervention, a high number compared to UK practice. A helpful indication of the time course of functional recovery post knee arthroplasty was also provided (Mizner *et al.*, (2005). Whilst subjects experienced worsening quadriceps muscle strength, functional test performance and range of knee joint motion at the one month post-operative time point, the average time taken to return to pre-operative levels was only two months post operation. Improvements were maintained at 12 months post-operation and, for stair climbing, were even equivalent to previously published data for healthy controls (Pettersen *et al.*, 2009). Overall therefore, post operative quadriceps muscle strengthening appears an important component of rehabilitation following arthroplasty.

In terms of late stage rehabilitation and return to sport, current practice again appears opinion rather than evidence lead; depending upon consensus recommendations based upon prosthesis wear, previous experience of activities, and joint loading (Kuster, 2002). Activities such as swimming, walking and golf are recommended/allowed whilst football and jogging are not recommended and no conclusion has been reached regarding downhill skiing (Kuster, 2002). A survey of 866 total knee arthroplasty patients found that, of the 253 who had been active in sport (generally low impact sport) preoperatively, 68 were unable to return to sport postoperatively because of their arthroplasty; pain and limitations in movement being the main reasons (Wylde *et al.*, 2008).

With regard to multidisciplinary rehabilitation, a recent Cochrane review of such programmes after hip and knee joint arthroplasty found no trials addressing outpatient programmes (Khan *et al.*, 2008); the review included one poor quality study comparing a 6

visit multidisciplinary care programme with a programme of up to 45 visits with no significant differences in self reported outcome.

## **2.9 CHAPTER SUMMARY.**

Chapter two summarised the background literature regarding osteoarthritis of the knee, the treatment and management of osteoarthritis of the knee, the prevalence of total knee arthroplasty and rehabilitation and outcome following arthroplasty. It can be seen that, despite being a long known and common condition, many questions remain concerning the symptoms of the condition and how to effectively rehabilitate patients following joint replacement.

# **3 CHAPTER THREE. EFFECTIVENESS OF PHYSIOTHERAPY EXERCISE FOLLOWING KNEE ARTHROPLASTY FOR OSTEOARTHRITIS: A SYSTEMATIC REVIEW OF RANDOMISED CONTROLLED TRIALS.**

## **3.1 CHAPTER OVERVIEW.**

This chapter will present the systematic review used to evaluate the effectiveness of post discharge physiotherapy exercise on function, walking, range of motion, quality of life and muscle strength, for osteoarthritic patients following elective primary total knee arthroplasty. The chapter begins with a background literature review, plus justification for the choice of methodology, before presenting the review protocol. The results will then be presented which will be followed by a discussion of the main findings, the strengths and weaknesses of the review procedures and the clinical implications from the systematic review. Finally a summary of the chapter is presented.

## **3.2 SYSTEMATIC REVIEW: INTRODUCTION AND BACKGROUND.**

### **3.2.1 SECTION OVERVIEW.**

This section will define and describe systematic reviews. It will outline the advantages for the project of performing a systematic review, explain the limitations of this approach and justify

the inclusion of a systematic review in this thesis. The section will particularly focus upon systematic reviews in relation to physiotherapy and the rehabilitation setting.

### **3.2.2 DEFINITION.**

Egger *et al.*, (2001a) combined several definitions to create the following:

A review that has been prepared using a systematic approach to minimising biases and random errors which is documented in a materials and methods section. A systematic review may, or may not, include a meta-analysis: a statistical analysis of the results from independent studies, which generally aims to produce a single estimate of a treatment effect.

Egger, Davey Smith and O'Rourke (2001) p. 5.

A systematic review should contain a clearly focused question (Main, 2003) and be conducted to an explicit and reproducible methodology (Greenhalgh, 1997) which strives to avoid the subjectivity and selection bias of traditional, or narrative, reviews (Main, 2003). The use of meta-analyses allows a single answer to a specific question to be produced from the multiple component studies (Main, 2003). Systematic reviews therefore are used to reduce or resolve uncertainty about evidence (Pettigrew, 2003).

### **3.2.3 BACKGROUND.**

The increasing body of published evidence makes it difficult for clinicians and researchers to remain up to date (Main, 2003) and for effective policies and guidelines to be developed by health service providers and decision makers (Mulrow, 1994). Previous research by the author of this thesis has found that physiotherapists in particular encounter problems locating and accessing literature (Beeston, *et al.*, 2001) with trials published in “almost every journal you can think of” many of which are not included in common databases such as Medline or Embase (Knipschild, 1994). Literature reviews are used to improve accessibility to research

findings. Traditionally narrative or journalistic reviews of the evidence have occurred, lacking a standardised and objective approach to the location and analysis of evidence and open to the personal biases of authors to present evidence supporting their own beliefs/theories (Greenhalgh, 1997). The consequences of this include confusion regarding choice of treatment, or poor/harmful practice being wrongfully continued (Chambers, 2001 p. xiv). The need for an objective, meaningful appraisal of the literature prior to undertaking research seems apparent. Systematic reviews can save resources and prevent unnecessary research occurring in areas where conclusive evidence exists (Mulrow, 1994). Reviews have been increasingly popular in medicine from the 1980s; especially after 1992 and the introduction of the Cochrane Collaboration (Montori *et al.*, 2003), an international organisation dedicated to the preparation and maintaining of systematic reviews regarding the effectiveness of health care interventions (Chalmers, 1993). The result has been that within a relatively short period of time systematic reviews have become “one of the cornerstones of evidence-based medicine” (Moher, 2008). As already mentioned, systematic may or may not include meta-analyses. Where meta-analyses are appropriate and possible for a review these are considered beneficial since they improve power<sup>1</sup> and precision<sup>2</sup>, answer questions not addressed by individual studies and address controversies arising from conflicting studies Deeks *et al.*, (2006). However, sometimes the wide scope of a review, or the diversity within component trials, mean that it becomes meaningless and potentially misleading to compare component studies (Egger, Davey Smith and O’Rourke, 2001a p.4-5; Main, 2003). In physiotherapy, it is often impossible to acquire data in a format permitting meta-analyses (Main, 2003) or sufficient high quality studies to achieve clinically relevant systematic reviews.

---

<sup>1</sup> Power is “ the chance of detecting a real difference as statistically significant if it exists” (Deeks *et al.*, 2006). Combining multiple studies may mean there is a greater chance of detecting an effect.

<sup>2</sup> Precision refers to the degree of variability of an observed sample statistic (Sim and Reid, 1999).



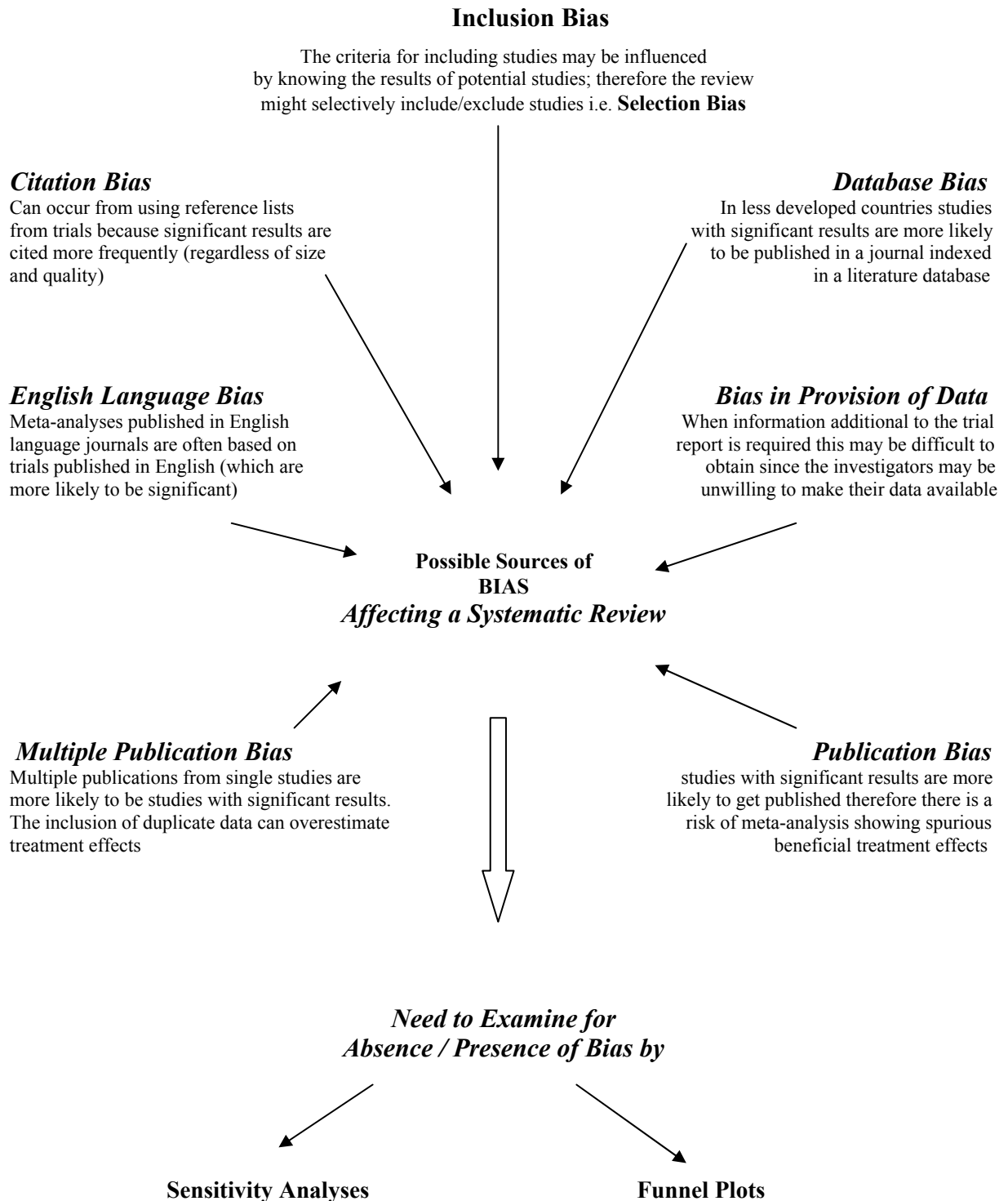
The use of systematic reviews, and meta-analyses in particular, remain controversial with some statisticians and clinicians not accepting either the mathematics or the findings: confusion caused by conflicting reviews upon the same topic and of reviews of small trials being subsequently contradicted by larger definitive trials has contributed to this dissent (Egger *et al.*, 2001c p. 43). These discrepancies may be due in part to poor methodologic quality in smaller trials (Kjaergard *et al.*, 2001) plus the results of a review may be unable to predict the results of large trials for up to a third of the time (Main, 2003).

It has been suggested that, for public health and social interventions, the “stainless steel” law of evaluation exists (Pettigrew, 2003), namely, “the more rigorous the review, the less evidence there will be that the intervention is effective” (Pettigrew, 2003). Furthermore, systematic reviews are able to answer single questions regarding specific interventions; they are less likely to provide clear answers to questions regarding complex interventions (Pettigrew, 2003) and it might be argued that this could be true for many areas of physiotherapy which incorporate complex interventions.

Despite being more explicit and objective than narrative reviews, systematic reviews may still be susceptible to bias (Main, 2003). An overview of sources of bias during the location and selection of trials has been provided by Egger and Davey Smith (1998) and summarised in the figure below. Such bias frequently results in an overestimation of a treatment. A further source of potential bias is reporting bias; there may rapid or delayed publication, or selective reporting, of a trial depending on the nature and direction of results (Egger *et al.*, 2001c p. 51-62).

Overall, despite the limitations of the approach, systematic reviews are advocated as the best available approach which currently exists to summarise and synthesise data (Main, 2003) and are considered a major advance in the objective review of evidence (Crombie and

McQuay, 1998). Please see the table on the next page for a summary of the advantages and limitations of this approach.



**Figure 2. Summary of the Sources of Bias when locating and selecting studies for inclusion into a Systematic Review (Summarised using Eggar and Davey Smith, 1998).**

---

## Advantages of Systematic Reviews

---

Vast amounts of literature can be summarised and presented to health care professionals and service decision makers

Promotes the use of focused questions

Review methodology is reproducible and explicit

Explicit methodologies may limit bias in locating and selecting studies

Increased reliability and accuracy of findings (i.e. effect sizes and risks)

Meta-analysis can increase the precision of results

Increased statistical power, useful for areas with low event rates or smaller effect sizes

The time taken to implement research findings into practice may be reduced

Can establish the generalisability and heterogeneity of results

Can generate new hypotheses for future research from main findings or sub group findings

May identify areas of inadequate research

Can prevent obsolete research being undertaken in an already adequately explored area

Efficient use of existing data

Physiotherapy studies with large numbers, and multi-centre trials, are still rare. Large trial results are often not yet available thus systematic reviews may be able to inform practice more quickly.

## Limitations of Systematic Reviews

---

Poor/ differences in methodological decision making can still lead to erroneous and/or conflicting results.

There is a lack of agreement regarding the assessment of study quality; many measures are inappropriate for physiotherapy studies.

Quality affects results. Inadequate randomization, allocation concealment and double blinding effect treatment estimates.

There is a lack of consensus regarding the appropriateness of analysis strategies used

There are often insufficient physiotherapy studies to enable successful systematic reviews to occur

Physiotherapy papers often do not report the information required for successful meta-analysis

Physiotherapy studies may be too diverse to allow meaningful meta-analysis to occur

Combining data may disguise / oversimplify distinctions between studies

Bias, such as selection bias, publication bias and language bias, can still be introduced

The results of a systematic review may not predict trial results for up to a third of the time

Systematic reviews are often criticised as being unable to give specific guidance; often because the studies provide few outcome evaluations

There may be a “stainless steel” law in operation, that is, the more rigorous the review, the less evidence there is that the intervention is effective.

Reviews may overlook important clinical details in the component papers, thereby reducing their validity

---

**Table 1. Summary of the advantages and limitations of systematic reviews for Physiotherapy** (Summarised using Mulrow, 1994; Greenhalgh, 1997; Hopayian, 2001; Main, 2003; Pettigrew, 2003).

A systematic review of the literature evaluating the effectiveness of post operative physiotherapy exercise following elective joint replacement appears worthwhile prior to undertaking the proposed trial for the following reasons:

1. To systematically determine the extent of previous research; checking the value of the question and whether it has already been adequately addressed (Mulrow, 1994; Knipschild, 1994).
2. To inform the development of the proposed trial intervention using previous studies (Knipschild, 1994), assisting the establishment of its theoretical basis (MRC, 2000) and lessening the chance of mistakes (Knipschild, 1994).
3. To inform the choice of appropriate outcomes for the trial.
4. To assist other clinicians seeking easily accessible knowledge (Greenhalgh, 1997; Main, 2003) in this common area of practice.
5. To provide, if possible, information regarding treatment effect size and its consistency (Mulrow, 1994).
6. If the level of heterogeneity permits, to increase the power to detect treatment effects (Pettigrew, 2003).
7. To determine the level of generalisability of the findings; since the diversity of multiple studies provides an interpretive context wider than in any single study (Mulrow, 1994).
8. To identify and seek to explain inconsistencies and conflicts in the research/data (Mulrow, 1994).

Issues needing consideration to produce a high quality review are summarised in Table 1 and highlighted, where relevant, in the protocol.

### **3.2.4 SECTION SUMMARY.**

Although the science of systematic reviews is not accepted by all, the methodology remains the most systematic and rigorous means currently available by which existing evidence may be synthesized. As such, systematic reviews are considered appropriate prior to developing the proposed trial to assess its potential value and to synthesize the available evidence.

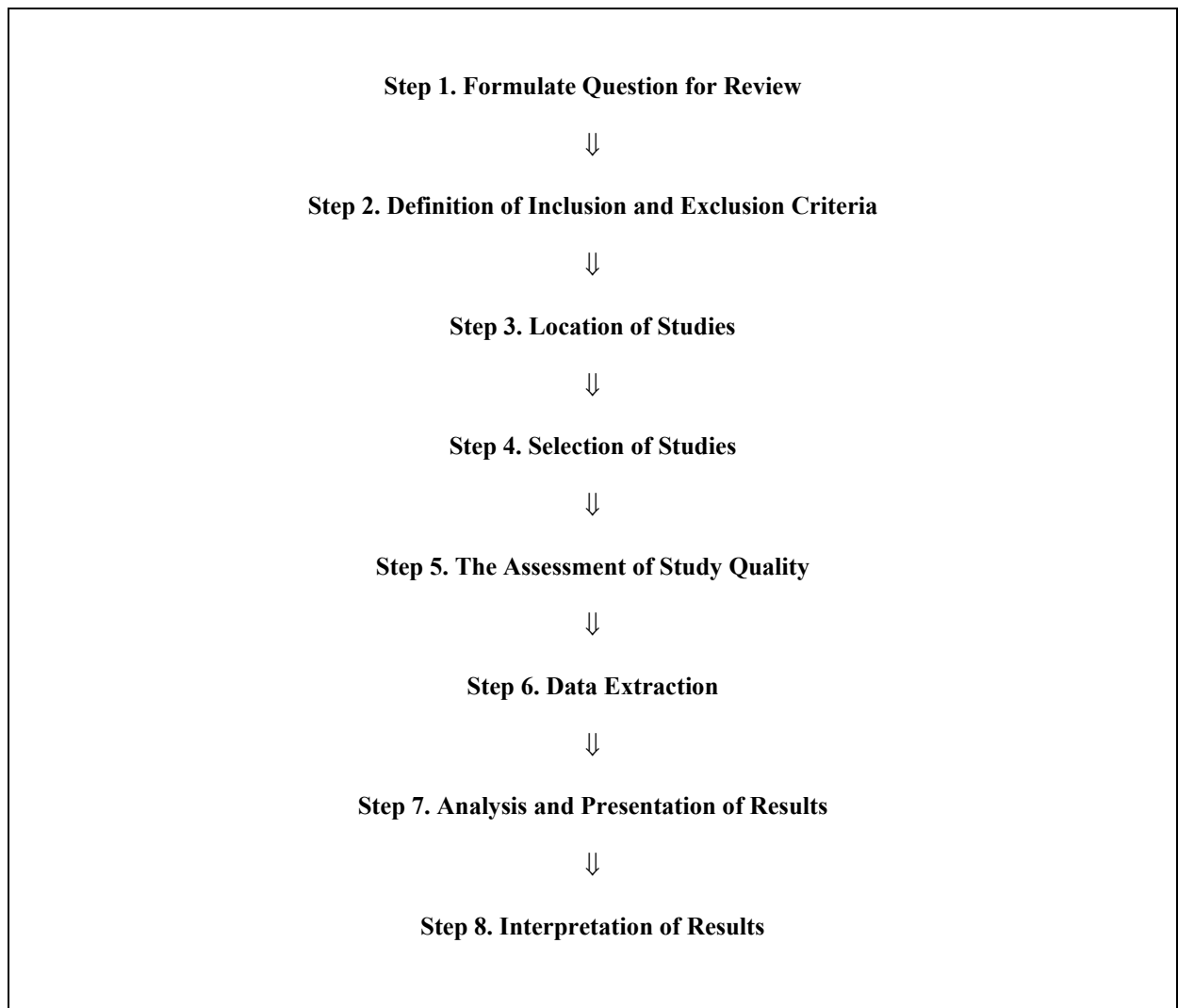
## **3.3 PROTOCOL.**

### **3.3.1 SECTION OVERVIEW.**

This section will present the eight stage protocol to evaluate the effectiveness of post operative physiotherapy exercise following discharge from hospital (in terms of improving function, walking, range of joint movement, muscle strength and quality of life) for osteoarthritic patients undergoing primary elective total knee joint arthroplasty. The protocol incorporates both a scoping review of available evidence, plus meta-analyses and meta-analytic summaries where appropriate, and follows the model outlined by Eggar and Davey Smith (2001 b, p. 24-28.) which is summarised the figure below. The Cochrane Handbook for Systematic Reviews of Intervention 4.2.5 (Higgins and Green, 2005a) was also used to assist protocol development. The Quality of Reporting of Meta-Analyses, currently known as the QUOROM statement although shortly due to be updated to the PRISMA checklist (Preferred Reporting Items for Systematic Reviews and Meta-Analyses), was also used throughout (Moher *et al.*, 1999; Moher, 2008).

### **3.3.2 Step 1. Formulate Question for Review.**

Question formation is an iterative process where the focus of the topic area is broadened and narrowed until a feasible and relevant question is defined (Montori *et al.*, 2003). This occurred during an initial literature search and during discussions at the ‘Systematic Reviews’ Module at the Department of Continuing Education at Oxford University in May 2005.



**Figure 3. The Eight Steps for the conduction of a systematic review (Adapted from Egger and Davey Smith, 2001b, p. 25).**

Although several narrative reviews were located (Brander and Mullarky, 2002; Kuster, 2002; Trudelle-Jackson, 2002) no systematic review of previous research was then available. The refined question then became the objective of this review as stated in chapter 1.

### **3.3.3 Step 2. Definition of Inclusion and Exclusion Criteria.**

The reviews included prospective comparative clinical trials relevant to the question. As mentioned in the introduction both knee and hip patients were included in this stage of the research and, since the searches were integrated, this section mentions both knee and hip patients. All located trials were included since the initial scoping literature reviews indicated few existing trials.

Participants: Patients undergoing elective primary unilateral total knee or hip arthroplasty for osteoarthritis who have received a physiotherapy exercise rehabilitation intervention following discharge from hospital post-operatively.

Intervention: The term physiotherapy exercise referred to any exercises or exercise programme advised or provided by physiotherapists/physical therapists during the rehabilitative period after discharge from hospital after surgery occurring in the out patient, community or home setting. The review is restricted to exercise rehabilitation, excluding other modalities such as electrotherapy which have been evaluated elsewhere.

Comparisons: Trials were included in the reviews if they:

- a) compared a physiotherapy intervention versus usual or standard care
- b) compared two different types of relevant physiotherapy intervention.

Outcomes: The choice of outcomes came from those used in trials located in the initial literature review; reassessed after data extraction and were the following:

- i) Self report measures of Function (in terms of level of activities of daily living)
- ii) Walking; the usual measure of mobility
- iii) Muscle strength
- iv) Range of joint motion
- v) Self report measures of quality of life .

### **3.3.4 Step 3. Location of Studies.**

Location was challenging due to:

1. Terminology: multiple terms describe physiotherapy and physiotherapy interventions; requiring more extensive and complex searching.
2. No single database holds all physiotherapy records (Knipschild, 1994; Main, 2003). Medline was insufficient (Dickersin *et al.*, 1994; Egger *et al.*, 1997).
3. The variety in outcomes increases search complexity since searching by outcome is more difficult to achieve.
4. Non English trials were included (Appendix II).

A broad area search was therefore undertaken, followed by more specific searches to check that as many records as possible were identified. It was accepted that this time intensive approach would identify many irrelevant papers and duplications but this approach appeared the most comprehensive method of locating the maximum possible number of trials present.

The strategy followed the procedures recommended by the Cochrane Handbook for Systematic Reviews of Interventions 4.2.5 (Higgins and Green, 2005b). The databases, searched in March 2005 and April 2007, were:

The Allied and Complementary Medicine Database (AMED) 1985-



Cumulative Index to Nursing and Allied Health Literature (CINAHL) 1982-

Excerpta Medica (EMBASE) 1974-

KingsFund 1979-

Index Medicus (MEDLINE)1966-

The Cochrane library (including Cochrane reviews, Cochrane Central Register of Controlled Trials, DARE)

Physiotherapy Evidence Database (PEDro)

The Department of Health National Research Register

Terms searched included:

Knee, hip, replacement, arthroplasty, trial, exercise, physiotherapy and physical therapy, occupational therapy, home programme.

Reference lists of all relevant trials identified by these literature searches were then checked for further trials. To check for recent and/or unpublished trials *Physiotherapy* (1985-April 2007 inclusive), *Physical Therapy* (1985-April 2007 inclusive), and *Journal of Bone and Joint Surgery (Britain) Conference Proceedings* (1985-2006 inclusive) were hand searched. Few additional records were identified; the team believed further searches unlikely to identify further quality trials and the search halted (Higgins and Green, 2005b).

#### **3.3.5 Step 4. Selection of Studies.**

Two researchers (the author and a supervisor, CS), one who knew the subject area well and one who did not, discussed the eligibility of identified studies to lessen any subjectivity regarding inclusion / exclusion decision making (Higgins and Green, 2005b). Any disagreements between the two researchers were to be settled by discussion with a third

reviewer. All clinical trials appearing to meet the review's inclusion criteria were considered eligible for inclusion.

### **3.3.6 Step 5. The Assessment of Study Quality.**

The quality of component trials is of crucial importance: if the "raw material" is flawed, then the findings of reviews of this material may also be compromised

Egger *et al.*, 2001c p. 45.

It is not easy to define trial quality (Jüni *et al.*, 2001a). Due to the numerous different quality checklists/ scales available, the model advocates assessing and presenting relevant methodological aspects for each paper instead, called the component approach. Jüni *et al.* (1999; 2001b) considered 25 quality scoring methods finding a wide variety in scales (from 3-34 items of varying dimensions and complexity) and component weighting. The same trial may score high or low depending on which scale is used; no single scale has been accepted as a "gold standard". Additionally, many scoring systems downgrade the quality rating of a trial if it is not double blinded. For many physiotherapy trials, such as those in this review, it is inevitable that patients and therapists know whether they are receiving the physiotherapy intervention or the control and this is not an indication of low/high trial quality. In my opinion this unfairly penalises such physiotherapy trials. The component approach, outlined by Egger and Davey Smith (2001, p. 24-28), addresses these problems. For this protocol key aspects recommended by Egger and Davey Smith (2001, p. 24-28) plus items were obtained from the CONSORT statement (Altman *et al.*, 2001) and the CASP guidelines (Critical Appraisal Skills Programme, accessed 14/05/05) were combined to provide a 22 item checklist (Appendix III). Similar analysis of individual quality components has previously been used in reviews of Physiotherapy (Shamley *et al.*, 2005).

The assessment of quality was undertaken independently by two reviewers (the author and a supervisor, KB). Due to the lack of consensus in the literature regarding whether reviewers should be masked to names of authors, journals, and institutions, sources of funding and acknowledgments and year of publication (see Appendix IV), the second reviewer was masked to these aspects but not the first reviewer, whose knowledge of the subject area made this process pointless.

Reviewer 1 (CML) obscured the papers for reviewer 2 (KB) by using a permanent black marker to delete the names of authors, journals and institutions, sources of funding, year of publication and acknowledgments throughout the text of full papers. This was done twice and then again after the articles were photocopied before the relevant details were truly obscured. In addition, (see Appendix IV), reviewer 2 was requested to complete a monitoring form (Appendix V) to determine the extent to which masking was achieved.

### **3.3.7 Step 6. Data Extraction.**

The checklist was reviewed and approved for use by KB and CS and piloted on one paper by the two data extractors, CML and KB, who then independently extracted all data and recorded this upon a form designed expressly for this purpose (Appendix VI). The checklist had 22 items. Items could be marked as yes, no, unclear or partial. Items were marked as yes if they fully and explicitly met the criteria laid out in the CONSORT standards (Altman *et al.*, 2001). CML also compiled the full intervention details of included trials to assist in the future development of the proposed trial intervention.

### **3.3.8 Step 7. Analysis and Presentation of Results.**

CML and KB discussed and compared the results of the data extraction. Any disagreements occurring were to be discussed with a third reviewer, CS. In the event this was not required. The two reviewers rarely totally disagreed, usually the disagreement was the more minor “yes” to “partial/unclear” or “no” to “partial/unclear” and 100% agreement was obtained upon discussion.

Where key study details were absent or unclear the authors were emailed and written to by CML requesting clarification or further information and the majority of authors replied with the sought information. All additional information received was also independently reviewed (by CML and KB) and discussed in the same way.

#### ***Summarising the data.***

The results of the quality assessment were presented in a standardised tabulated format to allow for easy comparison between studies. The results demonstrating the extent to which reviewer 2 (KB) was successfully masked to names of authors, journals and institutions, sources of funding and acknowledgments were calculated.

#### ***Inter-rater agreement.***

Percentage agreement between reviewers was calculated. Although useful as a quick guide and to assist in identifying whether a problem with the use of the kappa statistic has occurred (see below), percentage agreement is flawed because it does not account for some agreement occurring between the reviewers by chance or indicate where agreement/disagreement occurred (Altman, 1991 0. 404). The kappa statistic measures agreement whilst correcting for the amount of agreement that can be expected to occur by chance (Cohen, 1960). It has been defined as the ratio of the proportion of agreement (corrected for chance) divided by the

maximum number of times the raters could have agreed (also corrected for chance) (Yaffee, 2003). The kappa statistic does not allow for the degree of disagreement between reviewers plus a high interobserver agreement result can exist with a correspondingly low kappa value due to how the correction factor for chance works within the kappa statistic formula; part of what has come to be known as the kappa paradox (Feinstein and Cicchetti, 1990).

Due to the limitations of the kappa statistic, it is a controversial statistic and is not an unequivocal standard test to quantify agreement (Uebersax, 2002). The intraclass correlation coefficient (Landis and Koch, 1997), was therefore also used to inform the study about inter-rater agreement. ICC (2,1) was used because each subject was rated by each rater, with the raters being considered representative of a larger population of similar raters, and where a single measure was used. Unlike tests to solely measure association (such as Pearson's correlation coefficient for continuous variables) where scores could show high correlation but little actual agreement with each other, the intraclass correlation coefficient measures the magnitude of an effect and is used to provide a measure of inter-rater reliability (Howell, 2002). SPSS v12.0.1 was used throughout the review.

### ***Quantitative Data Analysis.***

Although variety was anticipated with regard to the participants, interventions and outcomes, the reviews were not anticipated to be so broad in scope that appropriate meta-analysis had the potential to be misleading or give an erroneous answer (Eggar *et al.*, 2001c p. 43) as might be the case if, for example, all physical treatment modalities (such as electrotherapy, osteopathy, exercise, acupuncture and so forth) were evaluated for pre and post operative joint replacement patients. Also, some diversity is useful since the generalisability and usefulness of meta-analyses is much improved if the component trials are not too narrow in scope

(Gøtzsche, 2000). In addition to narrative summaries quantitative analyses were undertaken. After the data was summarised, sufficient adequate data was obtained to allow the following:

### ***Meta-analyses.***

Meta-analyses were appropriate for knee function, walking, range of joint motion and quality of life. It was anticipated that the most frequent unit of analysis would be at the participant level, that is, the number of observations in the analysis would match the number of participants randomised. However, as the trials provided results for several follow up time periods, separate analyses were required for short (3-4 months) and long term (12 months) follow ups. Multiple intervention groups, which risk including some participants twice in the analysis (for example if control participants are included for treatment A versus control and treatment B versus control), were not encountered.

For each meta-analysis a summary statistic describing the treatment effect observed in each individual trial was calculated for each study and then the overall treatment effect was calculated as a weighted average of the summary statistics. The weightings reflect the amount of information each trial contains and was usually the inverse of the variance of the treatment effect, relating closely to sample size. Forest plots have been used to graphically display the results for each trial. In these plots the results of component studies are shown as squares, which centre on the point estimate of the results of each study, with a horizontal line running through the square demonstrating its confidence interval (Lewis and Clarke, 2001). The area of the squares represents the weight of the study within the meta-analysis (Egger and Davey Smith, 2001 p.30) The overall estimate from the meta-analysis are represented by a diamond shape at the bottom of the plot; the diamond's centre representing the pooled point estimate and the confidence intervals represented by the horizontal tips (Lewis and Clarke, 2001). Where the same measure was reported weighted mean differences were used, otherwise

standardised effect sizes were used (often labelled as small (0.2) medium (0.5) and large (0.8) (Cohen, 1969). Fixed effect models and 95% confidence intervals were used throughout. A fixed methods meta-analysis assumes that the true effect of treatment is the same value for each study and that therefore differences between study results arise by chance (Egger and Davey Smith, 2001 p. 34). In a random effects method, the assumption is that the individual treatment effects vary around an overall treatment effect (Egger and Davey Smith, 2001 p. 35) and these were not considered here as there was no compelling evidence of heterogeneity and estimating the between study variation is difficult with few studies. The differences were calculated so that positive differences indicate that the effect favoured treatment and negative that the effect favoured control/usual care. It is possible to perform sensitivity analyses where any studies of doubtful quality are included and then excluded in the meta-analysis and the results observed and discussed. If no change is observed it increases confidence in the results. Due to the small number of trials included in this review it was decided not to perform sensitivity and sub group analyses. Statistical analysis was performed using R 2.3.1 and the rmeta package (R Development Core Team, 2006) by Dr Michael Dewey. The data, data entry and selection of variables was provided and checked by Catherine Minns Lowe.

### ***Heterogeneity.***

An indication of the presence of statistical heterogeneity is present if the confidence intervals for individual trial results have poor / no overlap. Tests of heterogeneity, to assess whether there is greater variation between the results of trials than would arise by chance alone ( $\chi^2$ ), at a 5% significance level were performed but with the number of studies available in these meta-analyses it is accepted that these have low power. Since statistical heterogeneity is always likely to occur in a meta-analysis, instead of testing for the presence of heterogeneity,

some tests now quantify inconsistency so that the impact of heterogeneity on a meta-analysis can be assessed. One such commonly used test is the  $I^2$  (Higgins *et al.*, 2003), this was also calculated to give a measurement of the degree of heterogeneity between the trials in the meta-analysis. Statistical analysis was performed using R 2.3.1 and the rmeta package (R Development Core Team, 2005) by Dr Michael Dewey. The data, data entry and selection of variables was provided and checked by Catherine Minns Lowe.

### ***Bias.***

The assessment of publication bias, by the use of funnel plots, was felt to be inappropriate due to the small number of trials available for inclusion in the review.

### **3.3.9 Step 8. Interpretation of Results.**

The initial interpretation of the results of the review was undertaken by CML and, to ensure the review findings were appropriately placed into the surrounding clinical context, its accuracy and meaningfulness discussed with the other reviewers (KB, CS, MD). The review, plus the associated hip review, was then written up and submitted for publication (Minns Lowe *et al.*, 2007; Minns Lowe *et al.*, 2009).

## **3.4 RESULTS.**

### **3.4.1 SECTION OVERVIEW.**

This section will include the results of the selection and quality assessment processes, the characteristics of the studies included in the review, summaries of the interventions and the



quantitative data syntheses and analyses. The search strategies combined both knee and hip joint arthroplasty patients whilst the remaining results are for the knee only.

### **3.4.2 Location of Studies.**

The search strategy is summarised in the table below (full details in Appendix VII). 48 records were identified. As expected the trials identified were published in a wide range of journals (Knipschild, 1994).

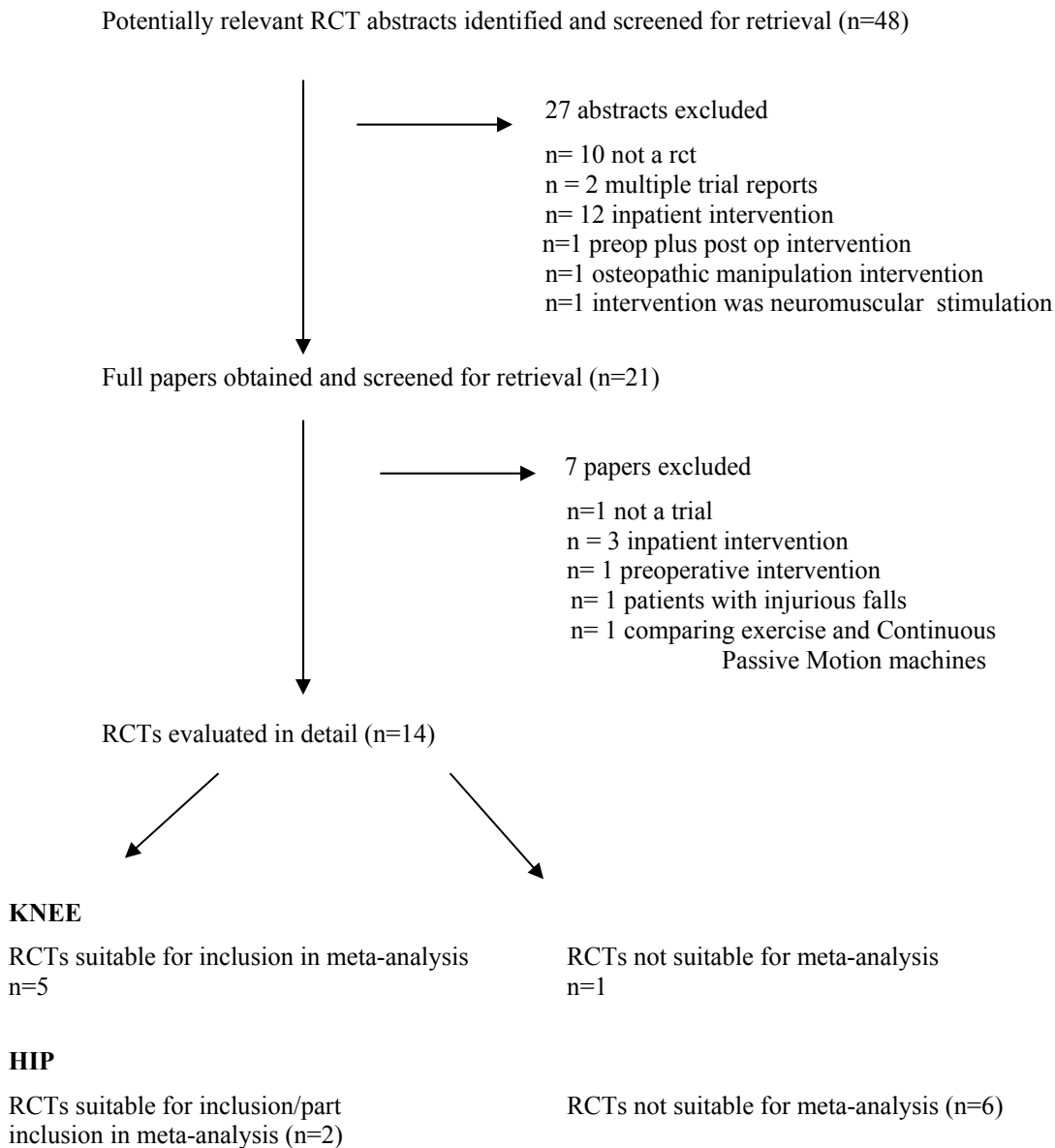
**Table 2. Search Strategy for Systematic Review.**

Source	Searches and Search Terms	Mar-Jul 2005 Hits* (no of new relevant records)	2005-Apr2007 Hits* (no of new relevant records)
KA24 :	1. "hip" OR "knee" (whole document)	587 (25)	180 (0)
AMED 1985 –	AND "replacement" OR "arthroplast\$" (whole document)		
CINAHL 1982-	AND "rehabilitation" AND "trial\$" (whole document)		
EMBASE 1974-	2. "hip" OR "knee" (whole document)	118 (11)	1 (0)
Kingsfund 1979-	AND "replacement" OR "arthroplast\$" (whole document)		
MEDLINE 1966-	AND "exercise" AND "trial\$" (title)		
	3. "hip" OR "knee" (whole document)	2 (0)	4 (0)
	AND "replacement" OR "arthroplast\$" (whole document)		
	AND "physiotherapy" AND "trial\$" (title)		
	4. "hip" OR "knee" (whole document)	39 (0)	14 (0)
	AND "replacement" OR "arthroplast\$" (whole document)		
	AND "physiotherapy" (title)		
	5. "hip" OR "knee" (whole document)	43 (8)	15 (0)
	AND "replacement" OR "arthroplast\$" (whole document)		
	AND "physical therapy" (title)		
	6. "hip" OR "knee" (whole document)	2 (0)	1 (0)
	AND "replacement" OR "arthroplast\$" (whole document)		
	AND "home programme" (title)		
	7. "hip" OR "knee" (whole document)	22 (2)	27 (0)
	AND "replacement" OR "arthroplast\$" (whole document)		
	AND "home programme" (whole document)		
	8. "hip" OR "knee" (whole document)	35 (0)	3 (0)
	AND "replacement" OR "arthroplast\$" (whole document)		
	AND "occupational therapy" (whole document)		
	9. "hip" OR "knee" (whole document)	0 (0)	3 (0)
	AND "occupational therapist\$" (title)		
Cochrane library:	1. Browsed by topic musculoskeletal		
Cochrane reviews	Search narrowed osteoarthritis	9	11
CCRCT	Search narrowed rehabilitation	80	18
DARE	2. General search term "joint replacement"		
PEDro	1. "joint replacement AND rehabilitation"	1 (0)	17 (0)
physiotherapy evidence database	2. "joint replacement"	5 (0)	45 (0)
Dept of Health	1. "joint replacement AND rehabilitation"	0	19 (1)
National	2. "joint replacement AND physiotherapy"	7	9 (0)
Research Register	3. "joint replacement AND exercise"	2 (0)	6 (0)
	4. "joint replacement AND physical therapy"	3 (0)	5 (0)
	5. "joint arthroplasty AND physiotherapy"	0	0
	6. "joint arthroplasty AND rehabilitation"	2 (1)	2 (0)
	7. "joint replacement AND occupational therapy"	5 (0)	5 (0)
Physiotherapy	Key journal - Hand search of contents pages	Nil new	Nil new
Physical Therapy	Key journal - Hand search of contents pages	Nil new	Nil new
JBJS [Br]	Hand search of all conference proceedings	2 new	1
Reference lists	Hand searching of papers included in the review	1 new	
Totals		965 (50)	386 (2)

\* Numbers following use of removal of duplicates commands when available.

### 3.4.3 Selection of Studies.

The 48 records were screened against the inclusion criteria first by CML and then by CML and CS at two meetings, 19<sup>th</sup> May 2005 and 14<sup>th</sup> July 2005. Consensus was reached regarding the inclusion / exclusion of all records (Appendix VIII) and summarised in the figure below.



**Figure 4. Flow diagram to summarise the selection of trials included in the reviews. This diagram is based on the Quorum statement by Moher et al., (1999).**

The following table gives details for the excluded studies. Following this decision making process, 14 records were considered to successfully meet the inclusion criteria for the systematic review and were selected for inclusion.

**Table 3. Studies Excluded from the knee systematic review.**

<b>Reason for Exclusion</b>	<b>Study</b>
Not a randomised clinical trial	Tum-Sugden 1976 Ulreich et al 1997 Benedetti et al 2003 Ritter et al 1989 Waters 1974
In patient intervention	Kolarz et al 1999 Beaupré et al 2001 Hewitt & Shakespeare 2001 Karst et al 1995 Kim & Moon 1995 Kumar et al 1996 Montgomery & Eliasson 1996 Hughes et al 1993 Lang 1998 Munin et al 1998
Osteopathic manipulation intervention	Licciardone et al 2004
Neuromuscular stimulation intervention	Stevens 2002
Pre operative intervention	Gursen & Ahrens 2003
Comparison of exercise and continuous passive motion	Worland et al 1998
Study halted early with no results	Stanley 2004
Duplicate trial report	Mockford et al 2006

#### **3.4.4 Pilot study: The assessment of study quality.**

The data extraction form (Appendix VI) was piloted on one paper and the results discussed by the two reviewers. Following the good level of agreement (26 out of 33, or 79%, categories achieving immediate agreement with 100% agreement after short discussion) between the two reviewers in this pilot phase, no changes were considered necessary to this form before use in the main review. Several issues were identified during extraction. In item 3 “does the study include the setting and locations where data was collected?”, the lack of agreement reflected

the success of blinding for location for the second reviewer. Lack of agreement was therefore desirable for this one item, which was subsequently removed from kappa, ICC and percentage agreement calculations carried out in the main review. High percentages of agreement were obtained for items where factual information was extracted from papers, for example, stating which outcomes were used and when extracting the results. Lower levels of agreement were extracted for items requiring a judgement or interpretation to be made regarding the quality of the data, for example items 20-22 relating to the quality of the discussion sections. Where discrepancies existed they tended to be between “Yes” and “Partial/Unclear”, “No” and “Partial /Unclear” and “Partial” and “Unclear” rather than between the larger gap of “Yes” and “No”. Discussion between the two reviewers easily resolved these minor differences in opinion.

#### **3.4.5 Main review: the assessment of study quality.**

All trials underwent full quality assessment except for one trial only available in abstract form (Mockford & Beverland 2004). For this review, studies were considered to be of good quality if they were deemed sufficiently robust to appropriately include in meta-analyses. The table below presents quality assessment findings for each study. One study (Codine *et al.*, 2004) was excluded from the meta-analysis because participants were allocated by alternation. All trial outcomes were measured by assessors masked to allocation. As the table shows, most studies clearly reported the flow of participants through the trial, justified study sample size and included intention to treat analyses. Several quality indicators were not fully discussed in all papers, such as allocation concealment and details regarding the implementation of randomisation methods.

**Table 4. Quality component checklist and quality evaluation of trials included in the review (n=6).**

Does the study/author information adequately contain the following:	Codine <i>et al.</i> , 2004	Frost <i>et al.</i> , 2002	Kramer <i>et al.</i> , 2003	Mockford & Beverland 2004 *see note 4	Moffet <i>et al.</i> , 2004	Rajan <i>et al.</i> , 2004
Rationale for study	Y	Y	Y	Y	Y	P
Eligibility criteria	Y	Y	Y	Y	Y	Y
Recruitment method	N	Y	N	Y	Y	P
Settings and location of study	N	Y	Y	Y	Y	Y
Intervention	P	Y	Y	Y	P* see note 5	N
Objectives / hypotheses	P	Y	Y	Y	P	Y
Defined outcome measures	Y	Y	N	Y	Y	P
Quality enhancers (e.g. multiple observations)	N	Y	P		N	N
Sample size determination	N	P * see note 3	N	Y	Y	Y
Randomisation	* see note 1	Y	Y	Y	Y	Y
Randomisation sequence generation	N	Y	N	Y	Y	Y
Allocation concealment	N	Y	N	Y	Y	N
Randomisation implementation methods	N	N	N	Y	Y	N
Blinding - participant	I	I	I	I	I	N
Blinding – of those administering the intervention	I	I	P	I	I	N
Blinding – outcome/assessments	Y	Y	Y	Y	Y	Y
Statistical methods	*see note 2	Y	P	Y	Y	Y
Flow of participants through each stage	N	Y	Y	Y	Y	Y
Recruitment and follow up dates	N	P	N		Y	Y
Baseline demographics	P	Y	Y	P	Y	P
Numbers analysed (and ITT)	N	Y	U	Y	*see note 6	Y
Summary of Results	N	Y	Y	Y	Y	Y
Estimated effect sizes	N	Y	N		N	N
Precision	N	N	N		Y	Y
Results for each outcome	Y	Y	Y	Y	Y	Y
Ancillary analyses	N	Y	U		Y	N
Adverse events	N	Y	P		Y	P
Interpretation	P	Y	P		P	P
Generalisability	P	Y	P		Y	N
Results placed into context	P	Y	P		Y	N
<b>Quality. Is the study of sufficient quality to include in meta-analyses?</b>	N	Y	Y	Y	Y	Y

**Key:** Y = Yes, included in paper/information to meet CONSORT level standards. N = not provided in paper/information. I = considered inappropriate/impossible. P = Partially evident in paper/information. U = not fully provided /explained in paper/information and therefore unclear/ambiguous. **Note 1:** 1<sup>st</sup> participant was drawn randomly then alternatively assigned (this process was witnessed). **Note 2:** additional information from author stated the sample size was determined by a statistician who calculated n = 30 in each arm to be sufficient. **Note 3:** This was a feasibility trial which provided sample size calculations as part of the results section. **Note 4:** Published abstract and information from authors only therefore this is incomplete. **Note 5:** intervention described but little description of home exercise programme described to both groups. **Note 6:** ITT analysis intended and performed but a per protocol analysis was presented because the loss to follow up in the control group favoured the intervention gr

#### **3.4.6 Masking in knee systematic review:**

The masking rates for authors (80%), journals (20%), author affiliations (80%) and funding sources (80%) were considered successful bar journal of publication.

#### **3.4.7 Reviewer agreement in main review:**

The following levels of agreement between reviewers were obtained 69.09%, kappa 0.524, intraclass correlation coefficient (2,1) 0.488 (95% CI 0.3 to 0.626) (Appendix IX).

Any instances where the two reviewers were not in initial agreement regarding the quality of a study following independent review were discussed until consensus was reached. Major disagreement was rare, usually disagreement was the more minor “yes” to “partial/unclear” or “no” to “partial/unclear” and 100% agreement was obtained. A third reviewer (CS) was available in the event of consensus not being reached but in the event this was not required.

#### **3.4.8 Study Characteristics.**

The following table summarises the characteristics of the included studies and provides information regarding the participants, intervention, main outcomes, and conclusions reached by the authors.

**Table 5. Study Characteristics of the Trials Evaluated in the Knee Systematic Review (n=6). TKA denotes total knee arthroplasty.**

<b>Paper</b>	<b>Participants (sample size)</b>	<b>Intervention (time of intervention)</b>	<b>Main Outcome Measures</b>	<b>Results</b>
1. Codine et al., 2004	Unilateral TKRApatients (60)	Submaximal training of hamstring muscle using eccentric isokinetic strengthening versus control of usual care (Day 10-30 post operatively)	Range of motion. Isometric muscle force. Knee Society Clinical Rating Scale	Significant difference between the two groups for extension only; favouring the intervention.
2. Frost et al., 2002	Unilateral TKA for osteoarthritis (47)	Comparison of a functional exercise group versus a traditional exercise group (post discharge)	Leg extensor power. Walking speed. Pain during walking (measured using 1 item from the Oxford Knee Score). Knee flexion	Feasibility study. No statistically significant differences between the 2 groups; trends in favour of the functional group.
3. Kramer et al., 2003	Primary unilateral TKA (160) total knee arthroplasty	Home based exercise group versus individual clinical based treatment (generally beginning within 1 week)	Knee society clinical rating scale scores. WOMAC. SF-36. Walk test. Knee flexion.	No significant differences between the groups.
4. Mockford & Beverland 2004	TKA (150)	Outpatient physiotherapy versus control of usual care (3 weeks post discharge)	Range of motion	Significant difference in range of motion favouring intervention.
5. Moffet et al., 2004	Primary unilateral osteoarthritis TKA (77)	Functional rehabilitation sessions versus usual care (2 months following surgery)	Functional ability – Distance walked in 6 minutes. SF-36. WOMAC	Significant difference in walking distance for functional group. At 2 & 4 month follow up (but not 12) the functional group had less pain, stiffness and difficulty performing activities of daily living.
6. Rajan et al., 2004	Primary TKA for monoarticular arthrosis (120)	Outpatient physiotherapy versus no outpatient physiotherapy alone (post discharge)	Range of motion in degrees	No significant differences between the 2 groups.



### **3.4.9 Summary of the Interventions and Comparisons.**

With the exception of one trial (Rajan *et al.*, 2004) in depth details of the intervention and comparison groups were available from the papers and authors. These are summarised in the next table. The trial interventions were similar to each other in that they provided additional physiotherapy exercises/treatment following discharge after total knee replacement, often involving functional weight bearing exercise programmes. Few details regarding the intervention are provided by Rajan *et al.*, (2004). Most of the interventions included functional weight bearing exercises. Only one trial differed significantly from this approach, Codine *et al.*, (2004) investigated the effect of eccentric isokinetic muscle strengthening using a CYBEX dynamometer. Interventions were usually started within two weeks of discharge. Out-patient programmes generally lasted up to twelve weeks while home exercise programmes were recommended for up to one year or indefinitely in one case (Kramer *et al.*, 2003).

The comparison groups were mainly control groups where no additional outpatient physiotherapy was organised. Patients were expected to continue with the traditional home exercise programme, namely isometric strengthening and range of movement exercises plus gait training/re-education, provided to all patients during their in patient stay.

**Table 6. Summary of trial interventions and comparisons included in the knee replacement trials (n=6).**

	Early programme provided to all trial participants	Intervention details	Comparison group details
Codine et al, 2004	Knee mobilisation (continuous passive motion and manual), isometric muscle strengthening for all knee and leg muscle groups, "proprioceptive enhancement," walking exercises	Submaximal hamstrings muscle eccentric isokinetic strengthening (passive resist mode). Torque produced > half of torque measured during testing. CYBEX dynamometer training speed of 10 degrees/s. Range of motion during flexion was conducted in active assist mode. Programme from day 10-30 post total knee arthroscopy, CYBEX 5 mins/day, 5 days/week for 3 weeks	Nil else added
Frost et al, 2004	Gait re-education, mobilising, and strengthening exercises (including active knee flexion, straight leg raises, inner range quadriceps, isometric quadriceps)	Warm up: sitting knee flexions 10 repetitions. Chair rise: baseline number set and increased every alternate day up to 2 mins, then repeat up to 3 times/day;. Walk: 1 min normal pace, increase 30 s/day until up to 10 mins, repeat 2-3/day. Leg lifts on to step/thick book: Baseline set. Increase by 1 per day until 2 min duration. Then repeat 2-3/day	Long sitting: static quadriceps (5 s hold); straight leg raise (5 s hold), inner range quadriceps (5 s hold). Supine: knee flexion heel slides. Standing: knee flexion (5 s hold). Sitting: long arc quads (3 s hold)
Kramer et al, 2003	Performed 3 times/day until 12 week check-up, then $\geq$ once a day. Stage I supine: knee flexion and extension. Long sitting: autoassisted knee flexion, ankle dorsiflexion in kneeeextension with calf stretch. Supine: isometric knee extension, inner range quadriceps. Supported sitting: hamstrings stretch. Sitting: active knee extension. 10 repetitions. Stretch/holds for 5 s. Stage II—prone lying: quadriceps stretch. Standing: quadriceps stretch, soleus stretch, Achilles stretch, knee flexion (progress to ankle weights). Supine: straight leg raise. Sitting: resisted knee extension, resisted knee flexion, sit-stand-sit. 10 repetitions, stretch/holds for 5 s. Optional exercises: exercise bike; standing wall sits 5-10 repetitions; standing slow short squats 10 repetitions	Outpatient physiotherapy weeks 2-12 after surgery: up to two sessions/week, each session about 1 hour. Exercises could be added/modified, therapeutic modalities (ice, heat, ultrasound), joint mobilisations, or other measures as appropriate. Patients requested to complete common home exercise only twice on clinic session days	Physiotherapist phoned patient $\geq$ 1 during weeks 2-5 and 7-12 after surgery for queries, advice. Patients able to phone with queries
Mockford and Beverland 2004	Supine: active ankle dorsiflexion and plantarflexion (10 repetitions), isometric quadriceps and hamstrings (3 repetitions), straight leg raise (repetitions vary), physiotherapist assisted knee flexion (5 repetitions), heels slides (3 lots of 10 repetitions, 2 mins rest between each set), active knee extensions over roll/bar (3 sets of 10 repetitions, as able, 2 mins rest between each set). High sitting: proprioceptive neuromuscular facilitation. Hamstrings pulley with 2 kg weight (3 lots of 10 repetitions, 2 mins rest between each set). Gait re-education with crutches/sticks as appropriate. Stairs practice	9 outpatient physiotherapy sessions in 6 weeks (2 sessions in weeks 1, 2, and 3, and 1 session in weeks 4-6). Week 1—heelslides, isometric quadriceps and hamstrings, straight leg raise, active knee extension over bar/roll and hamstrings pulley—all as in early programme. Proprioceptive neuromuscular facilitation, physiological mobilisations for flexion and extension (3 sets of 10 repetitions with 2 mins rest between sets). Weeks 2-3—standing: weight shifts (10 repetitions), quarter squats (10 repetitions). Prone lying: autoassisted quadriceps stretch (10 repetitions). Sit-stand-sit (10 repetitions). Gait re-education. Weeks 4-6—proprioceptive work in parallel bars. Gait re-education. Standing: stepping over cones, wobble board, step-ups (10 repetitions). Exercise bike for 5 mins	No outpatient physiotherapy
Moffet et al, 2004	"Simple exercises" to regain lower limb strength (quadriceps, hamstrings, hip abductors, and hip extensors) and to increase knee range of motion. Advice about knee positioning, ice application, and gait retraining	12 sessions of outpatient physiotherapy in 6-8 weeks, 60-90 mins each at clinic visit starting 2 months after surgery. Plus individual home exercise programme. Warm up (5-10 mins): lower limb flexion/extension, alternate ankle dorsi flexion-plantarflexion, hamstrings stretch. Specific strength exercises (15 mins): isometric knee extension in 0° and 60° flexion at visits 1-2; isometric hamstrings 60° flexion at visits 3-6: concentric eccentric hip abductors against gravity at visits 1 4. Functional task oriented exercises (15-20 mins): get up-sit down at visits 1-6; knee extensor strengthening in standing with Theraband at visits 1-6; controlled bilateral knee flexion-extension in standing at visits 1-8; unilateral knee flexion to 90° in standing at visits 7-10; climbing on platform/stairs at visits 3-12; walk backwards on slope and/or laterally while crossing lower limbs at visits 3-12; walk in place, with large amplitude hip and knee flexion and upper limb movements at visits 9-12. Endurance exercises (5-20 mins): walk at visits 3-12, exercise bike at visits 4-12. Cool down (10 mins), slow walk, strength, ice	Usual care. 26% patients had home visit. No attempt made to limit care. Information about usual care obtained by questionnaire and by phone interviews with patients and physiotherapist
Rajan et al, 2004	All patients given home exercise programme to follow on discharge.	Outpatient physiotherapy, 4-6 sessions	No organised physiotherapist

### **3.5 DATA SYNTHESIS.**

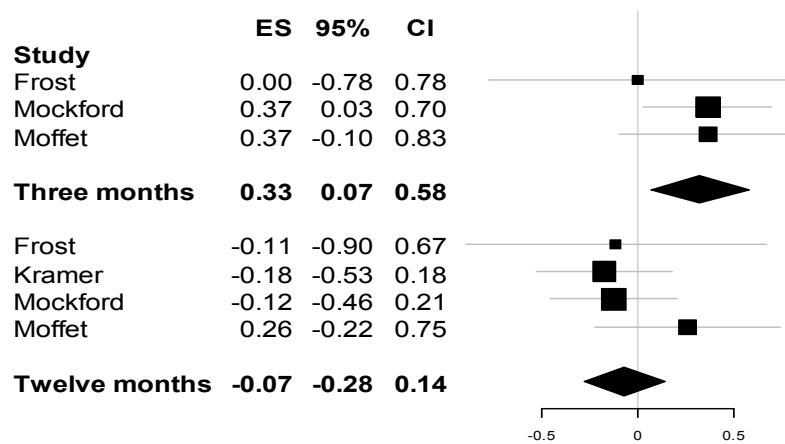
#### **3.5.1 Measures of function (5 trials, 494 participants).**

Five of the studies contained a measure of function (see study characteristics table). The measures used included:

1. The 12 item Oxford Knee Score (Mockford & Beverland, 2004) which measures functional ability, including pain, (scored 12-60, low score indicates high function). Frost et al (2002) used one item of this score.
2. The American Knee Society Clinical Rating Score (Codine *et al.*, 2004;Kramer *et al.*, 2003, Mockford & Beverland, 2004) measures pain, movement, stability and functional activity (scores 0-100, high score indicates favourable).
3. The 24 item Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) (Kramer *et al.*, 2003; Moffet *et al.*, 2004) has pain , stiffness and function domains scored as a percentage by Moffet *et al.*, and out of 0-170 for function by Kramer et al (low scores are favourable ).
- 4.The Bartlett patellar score (Mockford & Beverland, 2004) which measures anterior knee pain, quadriceps strength, and function (scores 3-30, high scores are favourable).

Within the individual trials, three demonstrated no observed significant differences between groups (1-3 study characteristics table). Frost *et al.*, (2002) showed significant within group differences for the treatment arm, indicating a treatment benefit. Mockford and Beverland (2004) presented no results in their published abstract but supplied summary statistics for their outcomes allowing us to include their study in the meta-analysis. Moffet *et al.*, (2004) demonstrated significant differences between the two groups, in favour of the intervention, at four and six months following arthroplasty, but not at 12 months.

The Forest plot below shows the three studies with data on functioning at three-four months and twelve months post surgery. Where studies included more than one measure of function it was decided to use the Oxford Knee and the WOMAC Scores since these encompassed all component trials. No trial included both the Oxford Knee Score and the WOMAC. At three-four months the standardised effect size is 0.33 (0.07 to 0.58) and the confidence interval does not include zero. This effect size is considered small to moderate (Cohen, 1969). At twelve months, with one additional study, the effect size is now close to zero at  $-0.07$  and the confidence interval includes zero.



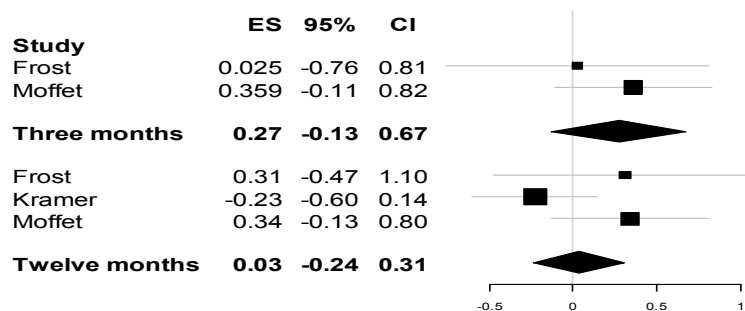
**Figure 5. Forest Plot showing Standardised Effect Sizes with Confidence Intervals for Function.**

### 3.5.2 Walking (3 trials, 284 participants).

Some form of walking outcome measurement was used in three knee replacement trials (2-3, study characteristics table). The measures reported included the following: walking speed over a 10m distance, measured in m/sec (Frost *et al.*, 2002) and a 6 minute timed walking test, measured in meters (Kramer *et al.*, 2003; Moffet *et al.*, 2004). The two timed walk tests varied in that Moffet *et al.*, report the use of a 50 m walkway.

The results from these trials are mixed. No significant differences were observed between groups in one trial (Kramer *et al.*, 2003), differences approaching significance were observed in another (Moffet *et al.*, 2004) trial while in the third trial significant differences within interventions groups were observed (Frost *et al.*, 2002).

The Forest plot below, using standardised effect sizes, shows the intervention to have no overall influence on walking at either three or twelve months.



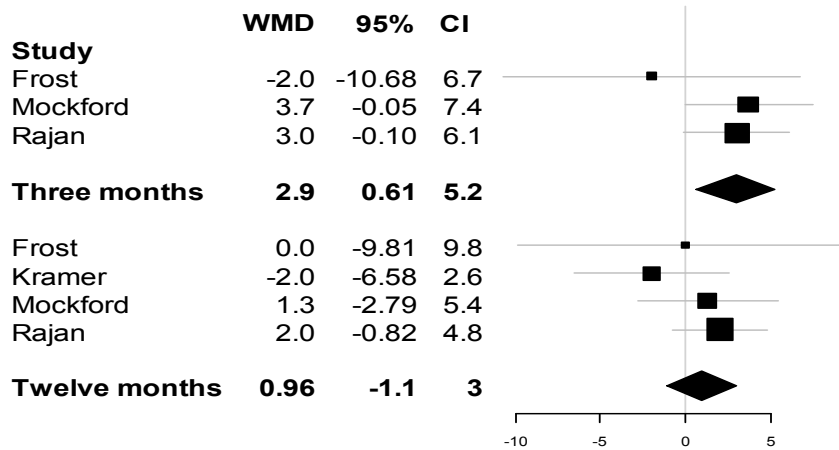
**Figure 6. Forest Plot showing Standardised Effect Sizes with Confidence Intervals for Walking.**

### **3.5.3 Range of joint motion (5 trials, 537 participants).**

Range of knee joint motion was used as an outcome measure for five of the total knee replacement trials (1-4,6, study characteristics table). Although all measurements were provided in degrees, the method of achieving results varied. Codine *et al.*, (2004) used a goniometer integrated into a dynamometer to measure knee flexion and extension whilst Mockford and Beverland (2004) measured active and passive flexion and extension using a goniometer. Frost *et al.*, (2003) and Kramer *et al.*, (2003) both measured active flexion only with Rajan *et al.*, (2004) providing findings for a single knee range of motion.

Once again, the results were mixed. Codine *et al.*, (2004) demonstrated a significant difference in knee extension between the two groups at ten days, however, despite randomisation, the two groups appeared by chance to have different extension at baseline. Mockford and Beverland (2004) concluded that there was a significant difference in active knee movement, in favour of the intervention group, but not in passive range. In the pilot study by Frost *et al.*, (2002) there was a trend that loss of range was less in the functional group than the traditional exercise group but no statistical significant difference was observed in the small study. In another two studies (Kramer *et al.*, 2003; Rajan *et al.*, 2004) no significant differences were also observed.

All the knee joint range of movement studies used the same measure, degrees, therefore the following Forest plot shows the weighted mean differences (WMD) and confidence intervals. The three month summary shows an increase of 2.9 degrees (0.61 to 5.2) with a confidence interval which does not include zero. This is considered a small to moderate WMD. At twelve months the effect is smaller, about one degree, and the confidence interval includes zero at this time point.



**Figure 7. Forest Plot showing Weighted Mean Differences with Confidence Intervals for Range of Motion.**

### **3.5.4 Quality of life (3 trials, 387 participants ).**

Quality of life measures were included in three knee trials (3-5, study characteristics table).

The SF-36 Health Survey provides an 8-scale profile of functional health and well-being

scores health dimensions with low scores indicating poor health. The SF-36 was used by

Kramer *et al.*, (2003) with the French translation of the same score being used by Moffet *et*

*al.*, (2004). The latter additionally provided the short form SF-12 Physical Component and

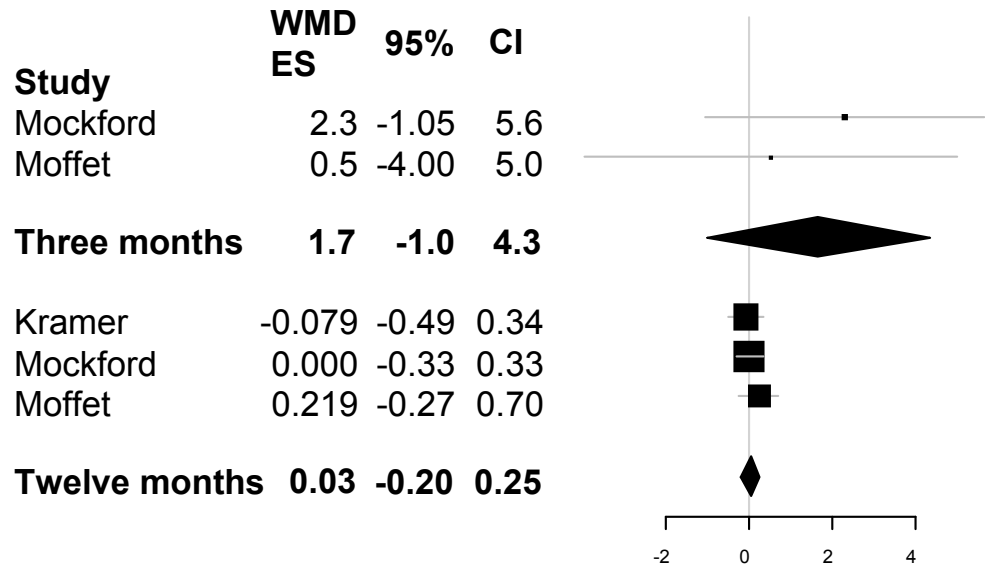
Mental Component Scores, as did Mockford and Beverland.

No significant differences between the groups were demonstrated by Kramer *et al.* (2003). Statistical analyses for this measure were not presented by Mockford & Beverland (2004). Moffet *et al.*, (2004) report small significant differences in favour of the intervention group were present for SF-36 role-physical dimension and the Physical and Mental Component scores at six month follow up but not at twelve month follow up.

The following Forest plot presents the studies with data provided regarding quality of life. At three-four months the studies used the same measure, the SF-12, and so weighted mean difference results are presented. However, at twelve months post surgery not all studies used the same measure and therefore the analysis has reverted to the use of standardised effect sizes which are presented in the table.

At three-four months post surgery the weighted mean difference is 1.7 (-1 to 4.3) with the confidence interval including zero. Thus there is a small effect seen in favour of the intervention. At twelve months the effect is close to zero with a standardised effect size of 0.03 (-0.2 to 0.25).





**Figure 8. Forest Plot showing Weighted Mean Difference (3-4 months) and Standardised Effect Size (12 months) with Confidence Intervals for Quality of Life.**

### **3.5.5 Muscle strength.**

Muscle strength itself was not directly measured in any of the trials included in the knee review.

### **3.5.6 Heterogeneity.**

The next table provides the results of the analysis of heterogeneity. As can be seen no major problems are indicated but the results are limited by low power.

**Table 7. Heterogeneity  $\chi^2$  and  $I^2$  Test Results for Function, Walking, Joint Range of Motion and Quality of Life.**

<b>Outcome</b>	<b>3-4 months post surgery</b>	<b>12 months post surgery</b>
Function	$\chi^2 = 0.78$ (df=2 p=0.68) $I^2 = 0\%$	$\chi^2 = 2.35$ (df= 3 p = 0.50) $I^2 = 0\%$
Walking	$\chi^2 = 0.54$ (df=1 p=0.46) $I^2 = 0\%$	$\chi^2 = 0.41$ (df = 1 p=0.52) $I^2 = 0\%$
Joint range of motion	$\chi^2 = 1.46$ (df=2 p=0.48) $I^2 = 0\%$	$\chi^2 = 2.28$ (df = 3 p = 0.52) $I^2 = 0\%$
Quality of life	$\chi^2 = 0.41$ (df = 1 p = 0.52) $I^2 = 0\%$	$\chi^2 = 0.91$ (df = 2 p = 0.63) $I^2 = 0\%$

## **3.6 DISCUSSION OF THE SYSTEMATIC REVIEW.**

### **3.6.1 SECTION OVERVIEW.**

This section presents a summary of the main findings from the systematic review evaluating the effectiveness of post operative physiotherapy exercise following elective joint replacement. The strengths and weaknesses of the review procedures are then discussed followed by a discussion of the clinical implications from the review.

### **3.6.2 SUMMARY OF KEY FINDINGS.**

The systematic review provides support for the use of physiotherapy functional exercise interventions following discharge, over traditional home exercise and advice programmes, to obtain short term benefit following elective primary knee arthroplasty. A small to moderate standardised effect size, in favour of functional exercise, was seen for function at three-four months post operatively. Small to moderate weighted mean differences, in favour of functional exercise interventions, were seen for range of joint motion and quality of life at three-four months post operatively. It is interesting to note that post treatment benefits were not seen at one year follow-up.

### **3.6.3 STRENGTHS AND WEAKNESSES OF REVIEW PROCEDURES.**

#### **3.6.3.1 Location of trials.**

The search strategy is considered comprehensive and to have been successful in locating relevant trials for inclusion in the review. Physiotherapy literature remains a difficult area to search, with numerous bibliographic data bases and un-indexed journals (Knipschild, 1994),

and while every attempt was made to identify studies it is possible other studies exist. However, these reviews remain the most comprehensive to date. The success of the search strategy used is indicated by a systematic review simultaneously published in 2007 by Dauty *et al.* This narrative review for literature regarding physical training in rehabilitation before and after hip and knee arthroplasty was limited to Medline and Cochrane databases, was date limited (1996-2006), did not include physiotherapy/physical therapy as a search term and did not search other databases likely to contain allied health professional research records. Consequently the review identified two trials from our fourteen knee and hip trials, even though all would have been relevant. The time consuming approach utilised for our reviews appears vindicated and, until location of physiotherapy trials becomes easier, shortcuts cannot be successfully undertaken for these types of reviews. The two non English language trials identified in the literature search were professionally translated to allow full inclusion and to prevent the introduction of language bias (Grégoire *et al.*, 1995; Moher *et al.*, 1996). Exclusion of non English trials would have removed one quarter of the trials later included in the hip review. Although the methodological quality of non English trials tends to be poorer (Juni *et al.*, 2002; Egger *et al.*, 2003) the trial identified in this literature search by Nyberg & Kreuter (2002) demonstrated that this cannot be assumed since this trial was of higher quality than some included English language trials (Minns Lowe *et al.*, 2009). This supports the decision to include non English trials in this search strategy.

### **3.6.3.2 Data Extraction.**

There were no apparent major problems with the data extraction processes used in this review; these processes will now be discussed.

### ***Masking.***

Previous blinding rates have varied from 54-73 % (Appendix IV). As such, the blinding rates achieved in these reviews appear satisfactory except for journal of publication. The repeated attempts to obfuscate study details could not disguise the layout of the journal. One way by which this might be achieved in the future would be to electronically scan and then edit trial reports prior to the review (Berlin, 1997). This was considered during protocol development but, with no scanner, was beyond available resources. Since reviewers may judge the quality of a paper by the journal in which it is published (Appendix IV) it is accepted that this may have affected the rating provided by the masked reviewer.

### ***Independent Reviewer Ratings.***

The independent reviewers showed good percentage agreement with each other and moderate agreements when using Cohen's kappa (Altman 1991 p. 404) and the Intraclass Correlation Coefficient (Landis and Koch, 1997). This level of initial agreement was considered acceptable since interpretation of some checklist items, such as generalisability and overall evidence, can be subjective. Following discussion both reviewers were in full agreement for all items for all papers.

The kappa statistical test depends on the number of categories, in this case four. Higher kappa values are obtained in 2 X 2 tables and lower kappa values for tables with more categories (Altman, 1991 p. 409). If this study had used a Yes / No score then a higher kappa value would have been achieved, but at the expense of introducing error by asking reviewers to guess where to place unclear or inadequate data. This may also have lead to a higher missing data rate since reviewers may have been unwilling to commit themselves to a Yes / No opinion.

ICC (2,1): SPSS provides the option of working out whether the error involved in the calculation is systematic: an absolute agreement type ICC measurement checks for systematic error while consistency type ICC measurement does not. Both types were calculated with little difference seen (Appendix VII and VIII).

### **3.6.3.3 Trial Quality.**

In the knee review trial quality was good overall. Of the five adequately randomised studies included in the meta-analyses, the majority were adequately powered with adequate allocation concealment strategies and outcome measurements were obtained by assessors blinded to treatment allocation (Altman *et al.*, 2001). Yet, like the majority of physiotherapy trials (Main, 2003), studies are relatively small with 554 participants in the five trials included in the meta-analyses and 614 participants included overall in the review. The most frequently used outcomes were function, predominantly subjective measures of functional ability, and range of joint motion as an objective measure. However whilst range of joint motion is important, its usefulness as an outcome measure of physiotherapy rehabilitation interventions is limited since other factors, such as prosthetic design, preoperative knee motion and surgical technique, also influence post operative range of joint motion (Sultan *et al.*, 2003). Muscle strength was not directly measured in any trial, although leg extensor power was included in one (Frost *et al.*, 2002), instead objective measures like walking were used.

The CONSORT statement, which was used to assist the assessment of quality of component trials in this review, has recently been extended to improve reporting in pragmatic trials (Zwarenstein *et al.*, 2008). This extension, like this review, accepts that blinding may not be done / possible within trials. Whilst this new extension is welcomed it is not believed that the quality of this review has been significantly lessened by being completed prior to publication of the extension. The need to comprehensively describe both the intervention arm

and the comparator arm in pragmatic trials was also emphasised as an indicator of quality in trial reporting in this new CONSORT extension (Zwarenstein *et al.*, 2008). The direct approaching of lead authors for additional information regarding quality indicators and the kind assistance of authors of component trials has enabled us to achieve this in this systematic review of published trials and we are grateful for their assistance.

#### **3.6.3.4 Data Analysis.**

##### ***Meta-analysis.***

The studies included in the knee systematic review showed some variation, such as the exercise type, dose and timing (as shown in the table summarising trial interventions and comparisons). Some reviewers believe meta-analysis is only appropriate for homogenous, similar or comparable data (Eysenck, 1994). Other researchers believe that heterogenous results from different designs are important since it is not always known a priori which methodological details, such as dose and timing, are most appropriate (Shrier, 2005). Systematic reviews that are narrow in scope can be misleading (Gøtzsche, 2000) as well as overly broad ones; there is a tension between a narrowly defined question and providing a widely applicable answer (Montori *et al.*, 2003). In this review the question was intended to provide a broad review including all relevant trials. Considerable thought and discussion occurred regarding combinability and all the review team are satisfied that the separate trials have been meaningfully combined and excluded in the review appropriately (Sacks *et al.*, 1987). The narrative and meta-analytic approaches have successfully summarised the current evidence available and generated hypotheses for future investigation. Despite the science of

meta-analysis remaining a matter of debate (Moore & Jull, 2006; Eysenck, 1994<sup>3</sup>) their use here is considered to have been beneficial in usefully informing the development of the proposed trial.

### ***Heterogeneity.***

Heterogeneity of studies included in meta-analysis is unavoidable; the question is whether its extent seriously undermines the conclusions being drawn from the systematic review (Davies & Crombie, 2001). In the knee systematic review none of the eight analyses reveals major problems with heterogeneity from the  $X^2$  tests but these were limited by low power. The  $I^2$  results indicated no observed heterogeneity (Higgins *et al.*, 2003). Overall, therefore, it is not possible to explore possible sources of heterogeneity in depth in this review and this limits clinical understanding regarding which parameters such as dose and type of exercise, are more effective (Thompson, 1994).

### **3.6.3.5 Clinical Implications.**

Presently, given the reduction in length of hospital stay, compressed in-patient rehabilitation and the limitations of the available evidence, it would appear reasonable to refer patients for a short course of physiotherapy following discharge after knee replacement to provide short term patient benefit. Whilst range of motion may be limited as an outcome measure of physiotherapy, the small to moderate standardised effect size obtained for function, which favours the intervention, is considered clinically important. This reflects actual improvements in one or more important aspects of function reported by patients after

---

<sup>3</sup> According to Eysenck, problems include the following: regressions are often non linear, effects are often multivariate rather than univariate, non homogenous data, the grouping of different causal factors may lead to meaningless estimates of effect, coverage can be restricted, bad studies may be included and the theory- directed approach may obscure discrepancies.



receiving the treatment intervention. No treatment benefits were seen at one year. This evidence is not considered conclusive. The content of the intervention could be better designed and further tested. Interventions to date have largely consisted of exercise programmes and gait rehabilitation, mainly targeting impairment and helping patients recover from the effects of surgery rather than specifically targeting activity limitations or participation restrictions. From the wider field of rehabilitation as a whole however such task training appears highly relevant. It was noted in a previous systematic review, assessing physiotherapy on functional outcome following stroke, that effective studies contained focused exercise programmes within which the relevant functional tasks were directly trained (Van Peppen *et al.*, 2004). Whether a feasible physiotherapy intervention of this type, supplied following discharge, impacts upon patient's functional ability at one year post knee replacement would appear a valid question.

It was not possible to include pain as a main outcome in this review since the studies identified in this review did not tend to measure pain as a specific outcome but rather as a component within measures of function. The pain patients experience from osteoarthritis and its influence upon function and the performance of objective measures also needs to be considered. With the recent introduction of a new osteoarthritis pain measure (Hawker *et al.*, 2008) it is hoped that future trials will include pain as a specific outcome in its own right.

### **3.7 CHAPTER SUMMARY.**

This chapter described the undertaking of a systematic review exploring to what extent is post discharge physiotherapy exercise effective, in terms of improving functional activities of daily living, walking, range of joint motion, muscle strength and quality of life, for

osteoarthritic patients following primary unilateral elective total knee arthroplasty? The systematic review provides support for the use of physiotherapy functional exercise interventions following discharge, over traditional home exercise and advice programmes, to obtain short term benefit following elective primary knee arthroplasty. Benefits were not seen at one year post operation. The results of this systematic review are not conclusive; they indicate however that functional exercise interventions following discharge show promise for this patient group. Overall, the number of available studies, their size, and their quality does limit this review and prevents the findings from being conclusive. It was both disappointing and surprising after the literature searching stage of this review to discover that so few published trials exist for such common areas of physiotherapy practice as knee and hip post arthroplasty rehabilitation. The need for further physiotherapy research for this patient group is evident and appears justifiable and worthwhile.

# **4 CHAPTER FOUR. COMPARISON OF POST DISCHARGE PHYSIOTHERAPY VERSUS USUAL CARE FOLLOWING PRIMARY TOTAL KNEE ARTHROPLASTY FOR OSTEOARTHRITIS: A RANDOMISED CLINICAL TRIAL.**

## **4.1 CHAPTER OVERVIEW.**

This chapter will present the trial undertaken to compare the effectiveness of a post discharge physiotherapy intervention in improving patient self report and objective functional outcomes versus usual care for osteoarthritis patients undergoing primary total knee arthroplasty. The choice of methodological approach will be justified and the merits and limitations of this approach explored. The development and rationale of the trial intervention will be presented. Next, the protocol followed for the trial will be detailed. The results will then be reported which will be followed by a discussion of the main findings, the strengths and weaknesses of the procedures used in the trial and the clinical implications arising from the trial's findings. Finally a summary of the chapter will be presented.

## **4.2 PURPOSE OF TRIAL.**

The purpose of this research was to discover whether patients undergoing primary total knee arthroplasty for osteoarthritis significantly benefited from receiving two additional home visit sessions of functional physiotherapy following discharge from hospital. Benefit was measured primarily in terms of self reported function. Objective measures were included as secondary measures: the leg extensor press to measure leg extensor power, the timed walk test to

measure gait and the timed sit-to-stand test to measure functional strength. The aims and hypotheses for the trial were presented in Chapter One.

### **4.3 METHODOLOGICAL APPROACH.**

#### **4.3.1 SECTION OVERVIEW.**

Experimental methods such as randomised controlled trials are commonly perceived to be the ‘gold standard’ design of clinical studies (Black, 1996; Jones *et al.*, 1996; Abel and Koch, 1999). Other methodologies have consequently often been judged less worthy (Lilford *et al.*, 1995). This section will begin by defining the term randomised clinical trial. The strengths and advantages of a randomised clinical trial research methodology will be presented. An examination of the limitations of randomised clinical trials will follow.

#### **4.3.2 DEFINITIONS.**

The randomised controlled clinical trial has been defined as:

A prospective study comparing the effect and value of intervention(s) against a control in human beings

Friedman *et al.*, 1998 p. 2

Randomised controlled trials have been further divided into explanatory and pragmatic trials.

Explanatory trials are:

intended to provide causal understanding of the effects of the studied treatment under optimal conditions

van de Windt, 1997, p. 124

Pragmatic trials, however, generally measure the effectiveness of an intervention in routine clinical practice (Roland & Torgerson, 1998). According to Sibbald and Roland (1998) the key elements within a randomised controlled trial are as follows:

1. Random allocation to trial treatment groups.
2. Participants and trialists should remain unaware, where possible and appropriate, which intervention was given until study completion.
3. Apart from the intervention itself, all participants are treated identically.
4. Intention to treat analysis is performed, that is that patients are analysed within the group to which they have been allocated regardless of whether or not they received the intended intervention.
5. The analysis approach focuses on the estimation of the size of the difference in predetermined outcomes between the intervention groups.

#### **4.3.3 THE VALUE OF RANDOMISED CLINICAL TRIALS.**

The benefits of randomised trials are outlined below.

##### *Assessment of effectiveness.*

Randomised clinical trials promote the assessment and quantification of treatment / intervention effectiveness (McPherson, 1994). McPherson (1994) opines that decisions regarding health care services, efficacy, equity and cost cannot be made on a scientific, intelligently discussed, basis if health care interventions are not adequately evaluated. While other study designs can detect association between interventions and outcomes, only randomised trials can rule out the possibility that the association was caused by another factor linked to the intervention and the outcome (Sibbald and Roland, 1998).

Furthermore trials can provide some information in order to predict the likelihood of an individual patient responding to an intervention in a similar way to the findings of the trial, providing that the trial is designed to provide data on participant characteristics likely to effect outcome (Mant, 1999).

***The value of randomisation.***

Randomisation is performed to form prognostically comparable groups at baseline (Koes and Hoving, 1998); since researchers can rarely identify and measure every prognostic variable for an intervention randomisation is relied upon to evenly distribute known and unknown prognostic factors across the groups. When patients are randomised to one of two or more treatments the act of randomisation aims to prevent potential bias in the assignment of patients by introducing unpredictability (Kunz and Oxman, 1998). Each patient is assigned a treatment through the play of chance rather than by advance prediction.

The effects of randomisation were explored in a useful systematic review by Kunz and Oxman (1998) who examined randomised and non randomised clinical trials and trials using concealed random allocation versus trials using inadequate concealment to determine the relationship between randomisation and estimates of effect. The review discovered that failure to use random allocation and concealment of randomisation generally resulted in overestimates of treatment effect (up to 150% or more of the estimated effect) although in some studies this bias went in the other direction and reversed the direction of the effect (up to 90% of the estimated effect) or masked an effect. Kunz and Oxman's (1998) conclusion is a powerful one:

Failure to use adequately concealed random allocation can distort the apparent effects of care in either direction, causing the effects to seem either larger or smaller than they really are. The

size of these distortions can be as large or larger than the size of the effects that are to be detected.

The need to ensure adequate randomisation and adequate concealment in the proposed trial is therefore clear.

### ***Defence of results.***

Abel and Koch (1999) also suggest that randomisation assists researchers in the defence of positive results to possible sceptics since the methodology is held in such high repute. Primacy is given to evidence from randomised clinical trials, sometimes summarised as the “if it moves, randomise it” approach (Croft, 1998).

### ***Prestigious Design.***

As mentioned the randomised clinical trial is commonly perceived to be the “gold standard” for comparing and evaluating different treatments (Pringle and Churchill, 1995). The randomised trial is an established methodology (Oakley, 1998), occurring since the 1940s (Swales, 1998); trials are currently deemed highly credible, attractive to funding bodies (Brewin and Bradley, 1989), and publishable, since randomisation can serve as an indication that the research has been carefully considered and thoughtfully carried out (Abel and Koch, 1999).

### ***Evidence based practice.***

In addition to the value of a single study, all studies pertaining to a specific treatment or intervention can undergo systematic review or meta-analysis, as in the previous chapter, to provide an overview of current evidence (Mulrow, 1994; Knipschild, 1994). This can then be used in order to promote evidence based practice (Greenhalgh, 1997; Main 2003) and, as in this trial, to assist the development of interventions for future studies (Knipschild, 1994).

### ***Blinding.***

Blinding is also frequently considered a beneficial component of randomised clinical trials. This aspect of trial design is the subject of Chapter five and is described and discussed there in detail.

### ***Control / Placebo groups.***

A major design benefit is that all clinical trials must contain a control group against which the intervention group can be compared so that any differences in outcome may reasonably be attributed to the intervention under investigation (Freidman *et al.*, 1998 p. 2). In laboratory experiments it is possible to provide or withhold an intervention, however in clinical studies this is often inappropriate or unethical. Placebo<sup>4</sup> controls may be used, primarily to enable patient attitudes to the research study to be as similar as possible in both treatment and control groups to minimise bias (Pocock, 1983, p. 93), but most often a new intervention is compared to the most effective current treatment or usual care (Friedman *et al.*, 1998 p. 2). This may mean the comparison group receives no intervention at all (Friedman *et al.*, 1998 p. 2). This is not considered disadvantageous by pragmatic trialists: pragmatic trials help clinicians to decide between a new treatment and the best current treatment, therefore the introduction of clinician or patient biases due to the lack of a placebo are not considered detrimental but accepted as part of their overall response to treatment and included in the overall evaluation (Roland and Torgerson, 1998). As Roland and Torgerson state:

...the treatment response is the total difference between two treatments, including both treatment and associated placebo affects, as this will best reflect the likely clinical response in practice.

---

<sup>4</sup> A placebo is an inert substance/procedure appearing similar to the trial intervention (Pocock, 1983 p. 9)



#### **4.3.4 LIMITATIONS OF RANDOMISED CLINICAL TRIALS.**

This section will identify and discuss the limitations of the randomised clinical trial as a research methodology. A previous commentary has attempted to redress the denigration of non randomised clinical trial methodologies (Black, 1996). Whilst the debate between trials versus observational studies continues (Ioannidis *et al.*, 2001), during his argument Black usefully highlighted many of the limitations inherent in trials. This influential commentary, plus other sources, have been used to summarise limitations of randomised studies in the figure below and key concerns to this trial are subsequently discussed.

**Experimentation might be unnecessary**

- Dramatic results may suggest a trial is not necessary (not true for the proposed trial)

**Experimentation may be inappropriate**

- trials insufficiently large to measure adverse effects
- may be unable to evaluate interventions designed to prevent rare events
- in trials where long term outcomes are involved
- too expensive (funding obtained for proposed trial so not appropriate)
- where random allocation reduces the effectiveness of the intervention
- where no appropriate outcome tools exist (not applicable to the proposed trial)
- for poorly understood areas of knowledge where qualitative designs may be preferred.

**Limitations of Randomised studies**

**Effects of Randomisation**

**Experimentation may be impossible**

(not the case for the proposed trial)

- Reluctance / refusal of clinicians or consumers to participate in a trial
- Ethical objections
- Political and Legal obstacles
- Contamination
- Scale (too many health care interventions to trial)

**Experimentation may be inadequate**

- trial patients may be unrepresentative
- low recruitment rate
- treatments may be unrepresentative
- trial clinicians may be unrepresentative
- to prevent excessive bias.
- in certain sites (e.g primary care)

**Figure 9. A summary of imitations of randomised studies** (summarised from Black, 1996; Abel and Koch, 1999; McPherson, 1994; Peto and Baigent, 1998; Koes and Hoving, 1998).

#### **4.3.4.1 Experimentation may be inappropriate.**

##### ***Trial size and rare events.***

As osteoarthritis is so common, patient availability is not considered a problem for this trial. However, the size of the this early phase II trial may prove to be of concern: small intervention groups may put the comparability of prognostic factors at baseline at risk, since the groups may be too small to evenly distribute the prognostic variables across all study groups, or to allow the identification of any adverse effects caused by the trial interventions (Koes and Hoving, 1998). The data analyses in the trial should be able to provide detail for the former but not necessarily for the latter. However, since the trial intervention includes established treatment components, such as exercise, it is believed that the risk of new adverse events occurring is unlikely.

##### ***Long term outcomes.***

Randomised clinical trials are rarely able to evaluate long term outcomes of medical treatments even when the outcomes of interest reach well into the future (Black, 1996). Black aptly uses arthroplasty to illustrate this point. In this trial, follow up will continue to one year, the time by which most improvement after knee arthroplasty is seen with most improvement occurring within the first 3-6 months (Jones *et al.*, 2007).

##### ***The influence of randomisation.***

A randomised trial may also be inappropriate where the very act of random allocation itself significantly reduces the effectiveness of the intervention (Black, 1996). Black's commentary argues that this can arise when the effectiveness of the intervention is dependent upon the active participation of the subject which then depends on the subject's beliefs and preferences. This can result in a lack of any subsequent difference in outcome between comparison groups

thereby risking an underestimation of the benefits of the intervention. This seems to be a similar issue, relating to potential bias, as that discussed in the earlier section of placebo/control groups; the same argument, that any subsequent bias is not considered detrimental since it would reflect clinical practice, could be forwarded again in rebuttal.

### ***Effects of Randomisation.***

There is evidence to suggest the belief that patients cite their aversion to randomisation as a main reason for their refusal to enter a trial (Llewellyn-Thomas *et al.*, 1991) which is of concern to all researchers seeking as representative a sample as possible. It was hoped that, because the proposed trial aimed to assess an additional intervention on top of usual care (with no placebo/control group) that this would not prove to be a serious issue. Randomisation is controversial; the feasibility of performing randomised clinical trials, the appropriateness of the design, ethical issues and the overall value of randomised trials in obtaining generalisable results are all areas of debate (Abel and Koch, 1999). This debate is not helped by randomisation being imbued with powers it does not necessarily possess; the erroneous belief that a researcher ensures comparability between treatment groups merely via randomisation thus ignoring differences occurring between groups following randomisation and before outcome / response measurement, which can adversely affect this comparability (Abel and Koch, 1999). Treatment effects, random errors, and patient / clinician motivation are amongst factors that Abel and Koch list as causes for treatment group differences. Also a dilution of intervention effect may be caused when prognostically heterogenous, rather than homogenous, groups occur after randomisation (Koes and Hoving, 1998). Plus, randomisation is not the sole basis for tests of statistical significance or for causal inferences on treatment effects, or sufficient to guarantee blinding (Abel and Koch, 1999).

### ***Ethical Issues.***

Ethical issues may also make a trial impossible to conduct. McLean (1995) states that there is a general acceptance that a study which is not scientifically valid is unethical; such studies are an abuse to the patient group concerned and a waste of scarce research resources. The need to obtain ethical consent prior to performing a study, from a relevant Ethics Committee, is seen as a way to assess the value of the study in these terms and to protect the participants in studies from unacceptable harm. Often, trials which expose patients to an intervention which is believed to be inferior to current treatment are regarded as unethical (Sibbald and Roland, 1998).

In a more subtle issue, Lilford and Jackson (1995) raise the question of whether or not trials are unethical without the prerequisite of equipoise. Equipoise is defined as:

“the point where there is no preference between treatments, i.e. it is thought equally likely that treatment A or B will turn out to be superior”

Lilford and Jackson (1995)

Lilford and Jackson (1995) argue that, whilst for many trials equipoise is desirable, there are times when equipoise is unachievable, an example might be when a central authority decides that a treatment needs evaluation for reasons of public interest. They also argue that randomisation without equipoise may actually be desirable in some instances, for example where access to an intervention is limited through scarce resources.

### ***Contamination.***

Black (1996) also describes the possibility of contamination if clinicians involved in a trial have to provide more than one intervention arm since the way in which the arms are delivered to patients may be influenced from this. Whilst recognised as an issue the current trial was designed in such a way that this could not happen so was not a concern for this trial.

### ***Number and Complexity of Interventions.***

As a further issue, health service interventions in primary care are often far more complex than the interventions assessed in clinical trials (Mant, 1999; Black 1996). Assessing very complex interventions can make trials impossible (Black, 1996). And, even when a trial is possible, the generalisation of findings to patients with their own individual complexities is difficult (Mant, 1999). The integration of qualitative research into the randomised trial can assist understanding of complex interventions (Mant, 1999) but cannot fully address this limitation. The Medical Research Council's Framework for the Development and Evaluation of RCTs for Complex Interventions to Improve Health (MRC, 2000) was developed to assist trials evaluating complex interventions (see next section on the development of the trial intervention) but cannot fully address the problem.

#### **4.3.4.2 Experimentation may be inadequate.**

A lack of consideration of external validity is a frequent criticism of randomised clinical trials (Black 1996; Rothwell, 2005). This raises three main issues pertinent to the current trial.

### ***Participants and Recruitment.***

Participants recruited to a trial may be unrepresentative of the entire patient population (Rothwell, 2005). Narrow eligibility criteria may mean only a small proportion of patients from those normally treated in everyday practice are included. The recruitment rate of participants may be poor, with many eligible patients not recruited to the trial (due to patient preference for example, or lack of commitment by the recruitment centres) which introduces the potential for selection bias to occur. Whilst recruiting / recording consecutive patients is recommended (Pringle and Churchill, 1995) this may not always be practical in busy

consultations. External validity is also decreased if patients who drop out from the trial are systematically different from those who complete the trial (Ward *et al.*, 1999). Patients who are recruited into trials may be those with more severe symptoms or those with consulting patterns that differ from the majority of patients (Peto *et al.*, 1993). Other authors, such as Mant (1999), argue that the eligibility criteria of patients participating in a trial are also likely to lead to unrepresentativeness in trials; with participating patients being more likely to be healthier, younger and of a higher social status than those patients ineligible to enter the trial. The amount of exclusions to a trial can, in some trials, be so excessive (Mant cites an exclusion rate of over 90% for one trial) that the findings of the study cannot be generalised to the majority of the patient population.

The single location in the current trial is viewed as beneficial in maximising consecutive recruitment in this phase II, exploratory, trial (MRC, 2000); although the generalisability of findings will be limited since only one hospital Trust is involved. All those approached will be recorded to allow the recruitment rate to be determined and the exclusion criteria selected to exclude as few patients as possible. The demographic data from this trial can be compared against data from the National Joint Registry for England and Wales to assess the representativeness of the sample in terms of age and gender.

### ***Treatments.***

Secondly, the treatments / interventions that the patients receive within a trial may not be representative of the treatments that everyday practice patients receive (Black, 1996). Patients may receive a higher quality of care and attention simply by participating in a trial, receiving treatment from experienced staff with optimal resources. Or trial protocols can differ from usual clinical practice (Rothwell, 2005). In this trial, although testing a new intervention, the

usual care group follow the hospital Trust's integrated care pathway which has not been altered in any way.

### ***Clinicians.***

Thirdly, clinicians consenting to being involved in a trial may have different characteristics to those clinicians not involved which may limit the generalisability of the results. This is thought to be true for GPs (Pringle and Churchill, 1995) with GPs who are already interested in research and who are already acquainted with trial procedures being more likely to participate in research trials (van de Windt, 1997).

### ***Location.***

There is a danger that research may become overly pragmatic by becoming so defined on current local circumstances that theoretical underpinning of interventions may become neglected affecting once again, the generalisability of results (Bradley *et al.*, 1999). Zwarenstein *et al.*, (2008) helpfully suggest that explanatory and pragmatic trials should be viewed as a continuum rather than as a dichotomy and it does appear that a balance is required between explanatory and pragmatic approaches within the design of a trial. With either extreme, of either explanatory or pragmatism, generalisability of findings may become limited. What seems to be apparent is that external validity issues around a trial are, although often never reported in trials or assessed in systematic reviews, inherently important when assessing the value of either explanatory or pragmatic trials.

### ***Excessive Bias.***

The introduction of severe bias into a trial can render a trial inadequate and misleading (Altman *et al.*, 2001). This whole section has considered potential sources of bias. The CONSORT statement and the recent extension of the CONSORT statement for pragmatic



trials promote the identification and detailed reporting of trials and the minimisation of bias (Altman *et al.*, 2001; Zwarenstein *et al.*, 2008). These statements have been referred to throughout the trial and the trial includes a CONSORT flow diagram in the results section.

#### **4.3.5 HOW MUCH PRAGMATISM IS PERMISSIBLE?**

In all comparative trials there is a need to create a tight, precise, trial design since any “sloppiness” in trial design may obscure any differences which occur between the two interventions (Jones *et al.*, 1996). Trials can be criticised as being too pragmatic if the interventions are described as being totally at the discretion of the health professional and if no clear definitions of the interventions are provided (Koes and Hoving, 1998). Without standardising or defining the interventions precisely then the generalisability of the results of a trial cannot be adequately assessed. Therefore, the description/treatment algorithm/protocol for a trial must be precisely stated for each trial. In addition, since many practitioners and patients want to know if the average trial results can be used to predict an individual outcome, trials need to be designed which provide sufficient information about the characteristics of patients which are likely to effect their outcome from an intervention (Mant, 1999). The “balancing act” required when executing a pragmatic trial between meeting rigorous methodological standards and implementing a feasible, achievable trial is challenging (Jansen *et al.*, 2006). For musculoskeletal rehabilitation trials however, it is usually possible for researchers and clinicians to carefully define a package of possible modalities from which the clinician providing the intervention can select appropriate treatment for patients; adherence to this package can be checked by completing case report forms, which provide a breakdown of treatment times and modalities, or video (Mullis *et al.*, 2006). The defining and refining of

the intervention, prior to the trial, is considered key to the quality and success of the trial by the author.

This is the approach utilised in the randomised clinical trial in this chapter; it is a pragmatic trial, considered feasible for the NHS to consider implementing should the intervention prove to be effective, but care has been taken with the development and description of the intervention. This forms the basis of the next section.

#### **4.3.6 SECTION SUMMARY**

The randomised clinical trial is a valuable, rigorous tool in determining the efficacy of an intervention. But it cannot be assumed that the use of a trial design in itself guarantees research quality and the provision of valid and reliable results for use with patients. Clinical decision making in practice involves many more factors than effectiveness alone. Trials are valuable where they address an appropriate question, where all stages of the trial are conducted and reported to a high standard and where the limits of the trial are clearly described. A trial design may be unsuitable and inappropriate for some research questions, settings and patient groups. Whilst a pragmatic design is considered appropriate to measure the effectiveness of the rehabilitation intervention proposed in this trial, the intervention must be defined with sufficient precision to allow the generalisability of the results to be assessed. The trial must also provide sufficient information about the characteristics of patients likely to effect outcome to be of use to clinicians making clinical decisions about their patients when the trial is reported. Insufficient attention to these issues risks an unethical trial demeaning patient involvement and exposing participants to any risks of potential harm from being in the trial for no purpose.

#### **4.4 THE DEVELOPMENT OF THE TRIAL INTERVENTION.**

The most challenging part of evaluating a complex intervention – and the most frequent weakness in such trials – is defining the intervention

MRC, 2000

##### **4.4.1 SECTION OVERVIEW.**

This section explains the challenges and issues encountered during the decision making processes required to develop the intervention for the trial. It will justify and clarify the decisions made and make the limitations of the intervention explicit. The section uses three headings, based on those recommended by Wade (2005), necessary to describe rehabilitation programmes/interventions, namely, the need to describe the theoretical model/basis of the intervention, the specific actions and activities of the intervention and the organisation of the intervention.

##### **4.4.2 THE THEORETICAL MODEL/BASIS OF THE INTERVENTION.**

The majority of musculoskeletal physiotherapy trials reviewed during this research do not explicitly state their underpinning theoretical framework. This is interesting since new developments are occurring which are intended to promote the ease and ability with which this may be done. One such development is the International Classification of Functioning, Disability and Health, or ICF as it is known (World Health Organisation, 2002). This framework has the potential to provide a universal language with which to describe health and its related areas, within a framework designed to encompass how people live with any health

conditions they experience as well as disease (Jette, 2006). Although further work is needed to clarify the ICF's concept and category differentiation, to develop tools to measure the ICF domains satisfactorily (Jette, 2006) and to adequately represent non Western cultures (Xie *et al.*, 2006), the ICF is believed to be a promising development since it conceptualises a “complex, multidimensional interaction that does not follow any predictable linear progression” (Bornman, 2004). The ICF is based upon a biopsychosocial model of health (WHO, 2002) and accepts that health conditions can be reversible (Bornman, 2004).

The ICF has been used in the development of a model of rehabilitation (Wade, 2005) which in turn has been used here to facilitate the description of the trial intervention. This model is summarised in the figure below which has been adapted to include the current trial to illustrate how it fits within this model.

## A SUMMARY OF WADE'S MODEL FOR REHABILITATION

### **Stage 1. The problem is presented**

*Patient consents to be in trial*



### **Stage 2. Assessment**

Data collected to: Identify problems

Understand genesis of problems

Identify prognostic factors

Understand patient's wishes/expectations

*Participant is assessed*



### **Stage 3. Goal setting**

Negotiation to determine short, middle and long term goals

*This will include highlighting individual rehabilitation issues during patient treatment as well as including the trial intervention package goals*



### **Stage 4. Intervention**

I. To collect data &/or to provide support &/or to give treatment

*For this trial: to collect data and to give treatment (i.e. to act to alter the natural history of recovery post knee replacement)*

II. Intervention to be in 1 or more of ICF domains

Patient Domains: Pathology, Impairment, Activities, Participation,

Contexts: Personal context, choice

Environment Domains: Social context, physical context, temporal context

*For this trial the intervention will aim to impact upon the activities, participant and physical domains.*



### **Stage 5. Evaluation.**

Comparison of patient state against goals set

Identify if any resolvable problems remain.

Decide if further action is required and either discharge or return to stage 1 as appropriate.

*For the trial this stage incorporates the data analysis phase*

**Figure 10. An Adaptation of the rehabilitation model suggested by Wade (2005).** *Italics are used to highlight information regarding the proposed knee replacement trial.*

Using the ICF and Wade models, the theoretical underpinning for stage 4 of the trial can begin to be outlined. The ICF identifies three levels of functioning; at the body / body part level, at the whole person level and at the person in a social context level (WHO, 2002).

Varying levels of disability can occur at one or more of these levels. Of particular relevance to the trial are the following:

1. Impairments; problems in body function or structure,
2. Activity limitations; difficulties an individual may have in executing activities,
3. Participation restrictions; problems an individual may experience in involvement in life situations, (WHO, 2002).

Whilst surgery may alter the level of impairment for trial participants, the trial intervention primarily aims to impact upon the level of activity limitations experienced by knee joint replacement patients following discharge from surgery. As the trial intervention aims to improve function, the level of participant restrictions may also be reduced. In addition, it is expected that the intervention will include the progression of walking aids for some participants, changing the physical context within the environment within Wade's model of rehabilitation (Wade, 2005).

The ICF (WHO, 2002) also makes a distinction between performance and capacity; the former describes what an individual actually does -the lived experience- whilst the latter describes an individual's ability to execute an action/task. This important distinction recognises that whilst a person may be capable of an activity they may, for various reasons, not actually do it in daily life. The proposed trial therefore includes tools to measure both self reported measures of function (although the possibility of a person saying they do an activity whilst not, in fact, actually doing it is also recognised as a limitation) plus measures of capacity.

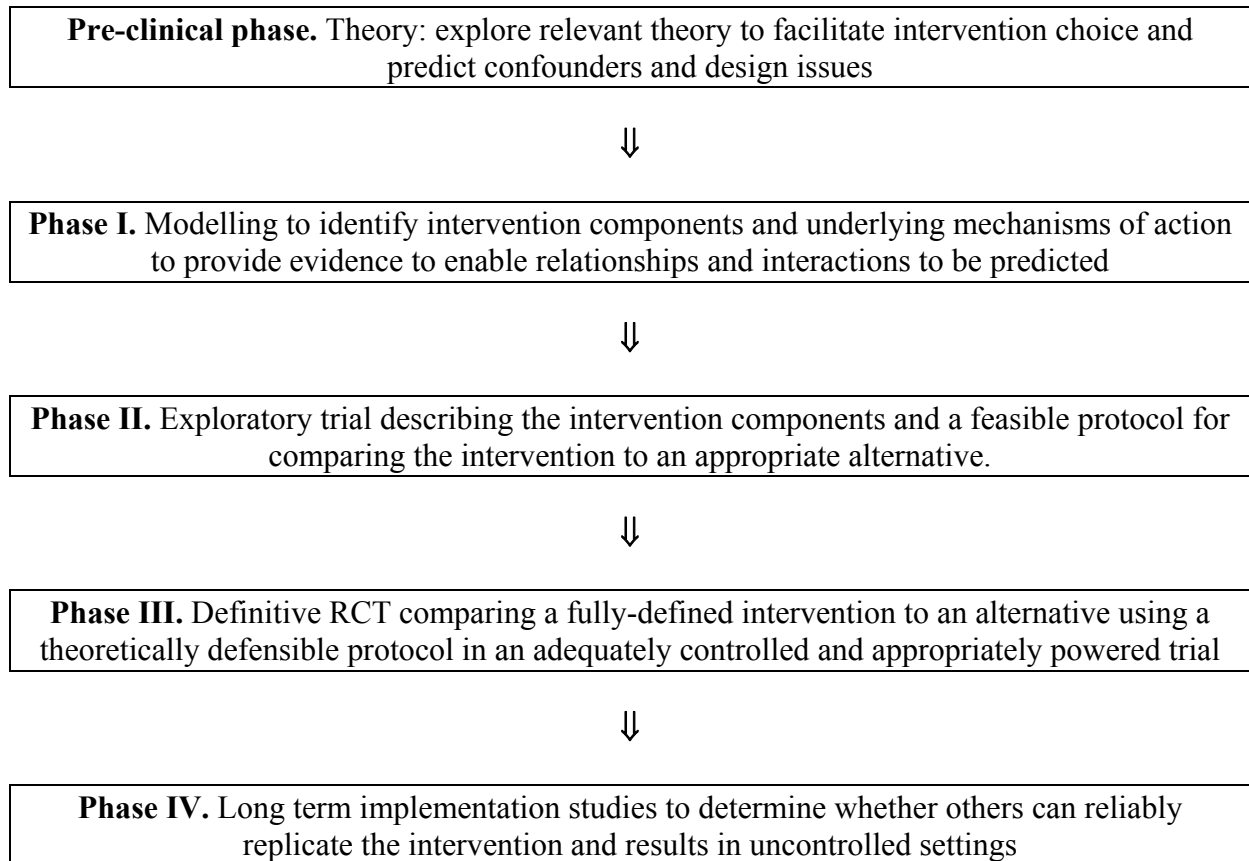
A further framework which is considered both valuable and pertinent for this trial is the Medical Research Council's Framework for the Development and Evaluation of RCTs for Complex Interventions to Improve Health (MRC, 2000). Complex interventions have been defined as follows:

Complex interventions in health care, whether therapeutic or preventative, comprise a number of separate elements which seem essential to the proper functioning of the intervention although the "active ingredient" of the intervention that is effective is difficult to specify

MRC, 2000

It has been said that "Rehabilitation is perhaps the archetypal complex intervention" (Wade, 2005) and certainly this trial meets the criteria for a complex trial intervention. In addition to the physiotherapy package itself are components such as the patient-physiotherapist interaction, the reassurance that post operative home visits may create, the home environment and specific advice the physiotherapists might offer. The MRC framework discusses all phases of work from early pre-clinical work right up to phase IV long term surveillance clinical trials. The framework is summarised below.

## A SUMMARY OF THE MRC FRAMEWORK FOR THE TRIALS OF COMPLEX INTERVENTIONS



**Figure 11. Summary of MRC Framework for Trials of Complex Interventions** (Adapted from the MRC, 2000).

The MRC framework is designed to provide a continuum of increasing evidence for complex interventions. The proposed trial will provide evidence for the early phases of this continuum for a new intervention package. It is not intended, or possible with one exploratory trial, to address phases III or IV. The pre-clinical phase uses all the available previous evidence and information to assist the identification of the study design and intervention type. For this study, a change in organisational structure which led to changes in the care provided to patients provided the initial impetus for this research.



The systematic review was used to identify intervention components likely to be clinically effective and to highlight important issues pertinent to the development of the intervention. Paper modelling of the proposed components identified a gap between the ICF and Wade models' and previous knee replacement interventions. Previous interventions mainly consist of home exercise programmes and gait rehabilitation and mainly target impairment. They do not usually include specific task training at the activity level. From the wider field of rehabilitation as a whole however task training appears highly relevant. It was noted in a previous systematic review, assessing physiotherapy on functional outcome following stroke, that effective studies contained focused exercise programmes within which the relevant functional tasks were directly trained (Van Peppen, 2004). In contrast, although studies focusing on reducing impairments (such as muscle strengthening or range of motion) did reduce such impairments, these changes did not lead to improved functional ability. Since previous knee replacement rehabilitation programmes have tended to target impairment rather than tasks/activities, an obvious question would be to ask whether the introduction of task/activity training would benefit musculoskeletal knee replacement patients as well as neurological stroke patients. It was therefore decided that task training should be included in the intervention to bridge this gap.

Although the trial intervention itself is new, its components are drawn from previous physiotherapy practice. Obtaining the opinions of physiotherapy clinicians and researchers to achieve a consensus opinion regarding the content of the intervention therefore seemed an appropriate way to add to the informal evidence base for the intervention as suggested by the MRC Framework. Additionally, the value of obtaining the views of participants' receiving the interventions was acknowledged and, in response, a future qualitative study was planned to explore participant's views regarding the intervention components and the content, level of

perceived value and acceptability of the intervention if it proved beneficial. As the MRC Framework (2000) opines, such qualitative research is useful in gaining understanding why something happens and in identifying the effective “active” components of an intervention.

The value of an exploratory trial is also discussed in the Framework (MRC, 2000). Such trials can test the feasibility of the trial design and intervention components, provide information for a larger trial regarding sample size calculations and trial monitoring procedures and also identify issues still to be addressed, such as intervention modifications or refining the intervention dosage. All these points are relevant to the proposed trial.

In summary, it therefore seemed possible to successfully integrate all three frameworks discussed in this section to theoretically underpin the development of the proposed trial. The figure below summarises how these frameworks have been amalgamated to achieve a theoretical basis for the processes used during the development of this trial intervention.

## **A SUMMARY OF STRATEGIES USED TO ACHIEVE A THEORETICAL BASIS FOR THE DEVELOPMENT OF THE TRIAL INTERVENTION**

**Pre-clinical phase.** Theory: explore relevant theory to facilitate intervention choice and predict confounders and design issues.  
A comprehensive literature review has been performed to highlight important issues for the intervention. The need for a systematic review of the intervention area was identified and undertaken.



**Phase I.** Modelling to identify intervention components and underlying mechanisms of action to provide evidence to enable relationships and interactions to be predicted.  
The systematic review of previous relevant trials was used to assist this modelling phase and to identify possible content for the intervention. The intervention is a treatment i.e. a package intended to act to alter the natural history of recovery post knee replacement. The intervention will aim to impact upon the activities, participant and physical ICF domains. Paper modelling, using flowcharts, was performed to model the intervention components and interrelationships. A consensus meeting with clinicians and researchers was held to agree a final intervention package.



**Phase II.** Exploratory trial describing the intervention components and a feasible protocol for comparing the intervention to an appropriate alternative.  
The trial will explore the effectiveness of the initial intervention. This is an exploratory trial exploring the feasibility of the intervention and to inform future work (for example by enabling the power of a main larger trial to be calculated). It is recognised that further phase II work in the future, for example to test alternative forms or 'doses' of the intervention would be required before a phase III trial is possible. A future qualitative study will explore participant's views regarding the intervention components and the content, level of perceived value and acceptability of the intervention.

**Figure 12. A Summary Of Strategies Used To Achieve A Theoretical Basis For The Development Of The Trial Intervention** (based upon the MRC Framework (2002) ICF (2002) and Wade's model of rehabilitation (2005).

### **4.4.3 THE SPECIFIC ACTIONS AND ACTIVITIES OF THE INTERVENTION.**

A pragmatic intervention package was planned with the therapeutic interaction occurring during each visit being accepted as part of the intervention. The intervention consisted of two

home visits. This number was decided upon as economically feasible for the NHS to implement should the trial show the short intervention to be effective. The first visit occurred within two weeks of discharge and the second 6-8 weeks post operatively. The timing of the start of the post discharge interventions was determined using the best available evidence. Interventions have begun within a week (Kramer *et al.*, 2003), 10 days post operatively, (Codine *et al.*, 2004), within three weeks of discharge (Mockford and Beverland, 2004) or 2 months postoperatively (Moffet *et al.*, 2004). Therefore an early intervention, within two weeks of discharge and three weeks post operatively was chosen. The 6-8 weeks time point was a pragmatic decision; patients routinely undergo a check at this time to make sure they are progressing satisfactorily and they commonly change walking aid around this time as well. In addition post operative pain relief often takes some time to occur postoperatively (6-12 weeks) which gives some time for pain to lessen. (Jones *et al.*, 2007).

During each visit the physiotherapist assessed the participant's function and progressed each participant's rehabilitation programme as appropriate. Objective reassessment of range of movement, muscle strength and observation of functional activities (including transfers, gait, posture and balance) enabled this progression of treatment. Gait re-education and progression/removal of walking aids occurred and a daily home exercise programme was provided and practiced with each patient.

The intervention package was defined using the systematic review (Minns Lowe *et al.*, 2007), the frameworks outlined in part one of this section above and a consensus approach incorporating the opinions of clinical, lay and research people. Consensus methods are appropriate "where unanimity of opinion does not exist owing to a lack of scientific evidence" since they allow the extent of agreement to be explored and disagreement resolved or identified (Jones and Hunter, 1995). An iterative process was used as follows:

1. Catherine Minns Lowe collated all available, relevant information regarding possible intervention components to be used to begin discussions during the consensus process.
2. This material was discussed with the PhD supervisors and revised.
3. The revised material was then discussed for an hour with the lay member associated with the trial. Issues such as dosage and the acceptability of possible content were discussed at this meeting. The material underwent further revision during and after this meeting which was subsequently approved by the lay member by post.
4. A consensus meeting of experienced senior staff was then held at a specialist orthopaedic centre. Approximately 25 senior staff attended this meeting. The staff groups represented included research staff, inpatient and outpatient musculoskeletal staff and rehabilitation staff from the Oxford Centre of Enablement. The format of the hour long meeting was as follows: a powerpoint presentation of the project to date was given, participants were given time to think about and rate possible components, full discussion of all possible elements and dosage parameters of the intervention then occurred, suggestions for revisions, improvements and additional content were invited and discussed. The intervention package, refined during this meeting, was rated and granted approval at the end of this meeting by the participants.
5. The refined intervention package, including all paperwork and exercise booklets, was then also discussed in three further separate meetings with a) research colleagues at a Physiotherapy Research Unit Meeting, b) the PhD

supervisors and c) the lay member. No further changes were suggested at this time and all meetings approved the refined version of the intervention package which was subsequently accepted as the final package.

The final package included exercises common to several programmes and considered practical to perform in patients' homes. The exercises have been reproduced in exercise booklets (See Appendix X) the repetition rate, frequency and progression of the exercises were left to the physiotherapist's discretion. Ideally, patients initially aimed for 10 repetitions of exercises 1-6 twice a day, building up to 20 repetitions and 5 repetitions of exercises 7-9 with 5 second "holds". Patients were asked to follow the programme for a minimum of three months. Again, these details were taken from the programmes used in the systematic review.

The functional exercises, in addition to walking, were follows:

1. weight shifts progressing to 1 – legged standing balance, progressing to eyes closed if appropriate,



**Illustration 6. Photograph of Exercise 1.**

2. partial knee bends/quarter squats,



**Illustration 7. Photograph of Exercise 2.**

3. Standing “wall sits”, knee flexion and extension leaning against wall,



**Illustration 8. Photograph of Exercise 3.**

4. bilateral heel raises progressing to unilateral heel raises,



**Illustration 9. Photograph of Exercise 4.**

5. knee raises with alternate arm raises,



**Illustration 10. Photograph of Exercise 5.**



6. step overs (alternate leading leg),



**Illustration 11. Photograph of Exercise 6.**

7. Achilles stretches in standing,



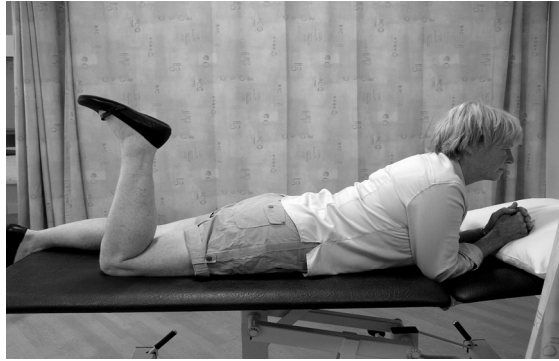
**Illustration 12. Photograph of Exercise 7.**

8. Soleus stretches in standing,



**Illustration 13. Photograph of Exercise 8.**

9. Quadriceps stretch in prone lying/standing (as able).



**Illustration 14. Photographs of Exercise 9.**

It should be noted that sit-to-stand and step ups were also commonly used exercises in the review. However these exercises are already included in the Usual Care patient information booklet used in the hospital at which the trial is based and which were provided to both groups. These exercises therefore were not included in the intervention package. Additional stretches were also included in the intervention package since only supported hamstrings stretches are included in the usual care patient booklet. Patients were taught how to stabilise their trunk and pelvis prior to lower limb movement and the quality of each movement was emphasized. Controlling the postural alignment of the trunk and pelvis is important since ideal postural alignment facilitates optimal lower limb movement and optimal muscle performance (Sarhmann, 2002 p 3-4, 367-368). Visit duration was timed.

The other trial arm received the usual care currently provided by the hospital with no additional input. Usual care included no routine physiotherapy organised post discharge. All trial participants followed the Trust's integrated care pathway and received the knee advice booklet used by the hospital to standardise the advice provided. These booklets were designed via multidisciplinary consensus, are used by all consultants and are in accordance with the

American Academy of Orthopaedic Surgeons guidelines. The advice and exercises contained in these booklets are essentially similar to those used as controls in previous trials. The exercises included non weight bearing exercises to regain range of movement plus isometric strengthening exercises and several exercises in weightbearing. The usual care exercises were as follows:

1. Static/isometric quadriceps contractions
2. Lying: Inner range quadriceps with rolled towel under knee
3. Lying: Knee flexion
4. Sitting: knee flexion and extension
5. Sit-to-stand
6. Placing operated foot onto a step and off again

#### **4.4.3.1 Repetitive task training.**

***Task training.*** In addition intervention group patients underwent the following task training:

Getting in/out of a car



**Illustration 15. Photograph of getting out of a car.**

Getting up from a chair at a table (similar to getting up after a meal)



**Illustration 16. Photograph of getting up from a chair at a table.**

Walking outside (e.g. similar to going to the shops and going out of the house)



**Illustration 17. Photograph of walking outside.**

## Stairs



**Illustration 18. Photograph of descending stairs.**

These tasks were identified and chosen using the six month post operative Oxford Knee Scores (OKS) provided by Dawson *et al.*, (1998) for each of scoring categories within each of the OKS functional questions. These scores showed that 48% of participants still had moderate/extreme difficulty with getting in/out of a car or using public transport; 31 percent found getting up after a meal moderate/extremely difficult/impossible; 53% found descending stairs moderate/extremely difficult/impossible; 43% found shopping moderate/extremely difficult/impossible. Since only 22% of patients reported similar difficulties washing and drying it was decided not to include training for this task in the intervention. Also, although, 87% of participants reported kneeling to be moderate/extremely difficult/impossible, some patients are specifically told to avoid kneeling following surgery and so kneeling was also excluded from the training package. However, kneeling could be practiced, as appropriate, as part of the individualised treatment provided to participants (see below).

#### **4.4.3.2 Individual Concerns.**

Finally, the physiotherapists answered/addressed any individual concerns from patients receiving the intervention. For example; if patients reported difficulties washing/dressing then these were practiced. These were noted to see if any common themes occurred.

All participants receiving the intervention were asked about activity limitations and participation restrictions so that these could be included in treatments and addressed where possible and appropriate.

### **4.5 TRIAL PROCEDURES.**

#### **4.5.1 ETHICS.**

Approval for this research project was obtained from Oxford Local Research Ethics Committee (Reference: A03.018). The main ethical issue presented to the committee was the need to ensure potential participants received sufficient time to consider participation prior to their attendance at the Pre Operative Assessment Clinic. A minimum of 24 hours notice was agreed. Appendix XI contains copies of the ethics approval and sponsorship and indemnity letters that pertain to this trial.

#### **4.5.2 DESIGN.**

A single blind randomised phase II controlled trial. The trial design incorporated the recommendations of the CONSORT statement (Moher *et al.*, 2001) and followed MRC guidelines for best practice.

#### **4.5.3 PARTICIPANTS AND SETTING.**

All patients undergoing an elective primary total knee arthroplasty for osteoarthritis residing within the community of Oxfordshire were eligible to participate providing they had none of the following exclusion criteria: patients undergoing bilateral arthroplasty, minimally invasive surgery and metal-to-metal implants, patients where further joint surgery was planned within the following twelve months, patients with inflammatory arthritis, patients whose existing co-morbidities prevented them from participating in the proposed treatment intervention and patients who were unable to provide informed consent. While the latter was an ethical issue all other criteria were excluded because of their influence upon outcome. No other restrictions, such as age or obesity were included to attempt to make the trial participants as representative as possible of patients undergoing knee replacement for osteoarthritis (Pocock, 1983 p. 35-38.). The trial was restricted to Oxfordshire to ensure reasonable travelling distances for the physiotherapist supplying the home based intervention.

#### **4.5.4 SAMPLE SIZE.**

As already mentioned, insufficient sample size may lead to misleading results (Peto and Baigent, 1998) but at the time of developing this early study there was little information available from which to calculate sample size. The sample size was based upon the use of the Oxford Knee Score as the primary outcome variable. 1000 simulations were performed to determine the sample size by the trial statistician Dr Michael Dewey. The simulations were carried out using a standard deviation of 10 from unpublished data from the original paper (Dawson *et al.*, 1998), obtained from the developers of the OKS (personal communication), autoregressive of order 1 correlations and allowed for the baseline and 3 follow up time



points. The developers of the OKS suggested a clinical difference of 2.5, one quarter of the standard deviation, would be reasonable given the lack of available evidence (personal communication). As will be described later the OKS is a self administered 12 item questionnaire whereby the patient rates, from 0-4, the difficulty and pain associated with performing activities of daily living, such as walking and putting on socks/tights. A score from 0-48 is obtained. A three point difference was chosen for this trial since this signifies that at least one or more functional activities have become easier and less painful to perform. For a three point clinical difference and power of 80%, n=40 completers for each arm were required. To allow for 10% drop out the total sample size was set at 88. The NOC expected approximately 576 primary knee replacements per year at the time of the study so recruitment numbers appeared feasible.

#### **4.5.5 BLINDING.**

Since it was not possible to mask the participants, or the physiotherapists providing the interventions, to treatment group allocation following randomisation the study is a single blind trial. However, all outcome measurements were provided by physiotherapists blinded to treatment allocation and participants were repeatedly requested not to mention which treatment they received during follow up outcome measurement appointments. Additionally, the physiotherapists performing the assessments were asked to record the extent to which masking was successful on each patient assessment form. This subject is the topic of chapter five.

#### **4.5.6 OUTCOME MEASURES.**

Copies of all outcomes can be seen in the trial questionnaires contained in Appendix XII.

##### **4.5.6.1 Primary Measure: The Oxford Knee Score (OKS).**

*Purpose of the Score and reason for its inclusion in the trial.* The Oxford Knee score was specifically designed and validated to measure function and pain following knee joint arthroplasty (Dawson *et al.*, 1998). The OKS is in widespread use, is a recommended measure for use in assessing outcomes in knee patients (Garratt *et al.*, 2004; Liow *et al.*, 2003) and has the advantage of being used in previous arthroplasty physiotherapy trials (Mockford and Beverland, 2004; Frost *et al.*, 2002).

*Completion.* The scores have been found to be easy to complete with an average completion time of 9.6 minutes by patients (Dunbar *et al.*, 2001). Response rates are usually high, ranging from 85.9 -90%; greater than for the WOMAC or Lequesne (Dunbar *et al.*, 2001; Robertsson and Dunbar, 2001; Whitehouse *et al.*, 2005).

*Scoring Approach.* The OKS is a short, practical self administered 12 item questionnaire whereby the patient rates, from 0-4, the difficulty and pain associated with performing activities of daily living, such as walking and putting on socks/tights (Dawson *et al.*, 1998). The total score is a summation of the 12 individual question ratings. Initially the score ranged from 12 (high level of function) to 60 (low level of function) but this was found to be counterintuitive and was revised (Murray *et al.*, 2007). The revised method has been used for this trial. Each item is scored from 0-4 and summated to give a score from 0-48, with 48 being the best outcome.

*Validity.* The score has been extensively validated against 10 measures (Garrett *et al.*, 2004). It was originally validated favourably against the American Knee Score, the Stanford Health

Assessment Questionnaire and SF36 (Dawson *et al.*, 1998) and since then favourably against the disease specific WOMAC and Lequesne Scores amongst others (Dunbar *et al.*, 2001). It has also been correlated with patient satisfaction (Spearman's non parametric correlation coefficient 0.68) (Robertsson and Dunbar, 2001).

*Reliability.* The Score demonstrates high internal consistency, with a preoperative Cronbach's alpha score of 0.87 and post operative score of 0.93, and high test-retest correlation  $r=0.92$  (Dawson *et al.*, 1998). High internal consistency scores have since been repeated, Cronbach's alphas of 0.93 (Dunbar *et al.*, 2001) and 0.92, leading some to believe that the scale includes redundancies (Whitehouse *et al.*, 2005). However, the quick completion time means this does not appear overly onerous for patients (Dunbar *et al.*, 2001). Intraclass correlation coefficient values range from .94 (Dunbar *et al.*, 2001).

*Specificity.* Like most knee scores, the coexistence of hip or spinal pathology can impact upon, and contribute to, OKS scores (Harcourt *et al.*, 2001) although the developers of the score believe the OKS to have greater specificity than other knee scores (Murray, *et al.*, 2007).

*Responsiveness to change.* A small study (six observers and 29 patients) by Liow *et al.*, (2003) found the OKS produced the greatest score ranges, normalised mean scores of 10.4-100, when compared against the American Knee Society Score, normalised mean scores of 17.1-97.2, and the British Orthopaedic Association score, normalised mean scores of 36.5-99.4. The OKS therefore was the most responsive to change. Dawson *et al.*, (1998) calculated an effect size from the pre and post op OKS scores, reporting an effect size of 2.19 and finding that change scores were significantly greater for those patients reporting the greatest improvements.

*Floor and Ceiling effects.* Dunbar *et al.*, (2001) determined the floor effect of the OKS to be 6.76 and the ceiling effect to be 0.1%. These scores were similar to the Lequesne score and superior to the WOMAC.

#### **4.5.6.2 Secondary Measures:**

#### **4.5.6.3 The Knee injury and Osteoarthritis Outcome Score (KOOS.)**

*Purpose of the Score and reason for its inclusion in the trial.* Measurement instruments rarely achieve universal acceptance and approval and critics of the OKS suggest that the scale lacks detail and should not be used where detailed assessment of outcome is required (Whitehouse *et al.*, 2005). It was therefore decided to include the KOOS, a more detailed knee function score, (Roos and Lohmander, 2003) to supplement the measurement of self reported knee function. The KOOS is a self report knee-specific instrument designed to evaluate both short and long term knee conditions, including osteoarthritis (Roos and Lohmander, 2003). Like the OKS, the KOOS is recommended as a measure to assessing outcome in knee patients (Garratt *et al.*, 2004).

*Scoring Approach.* The KOOS consists of 42 items in the five subscales of pain, other symptoms, function in daily living, function in sport and recreation and knee related quality of life. The Koos is scored 0-100 (100 indicating no symptoms and best outcome) for each subscale. A change of 8-10 points or more is considered a clinically significant change following anterior cruciate ligament surgery however this has not been explored with knee arthroplasty (Koos and Lohmander, 2003).

*Validity.* Content validity. When knee arthroplasty patients were asked to report the importance of the subscales when deciding to undergo surgery, 91-95% of patients felt all subscales were extremely or very important except for the sport and recreation subscale which only 515 believed to be extremely or very important (Roos and Toksvig-Larsen, 2003).

*Construct Validity.* The KOOS has been validated against the SF-36 score with knee arthroplasty patients showing moderate to strong correlations in the KOOS scales and SF-36 subscales designed to measure similar constructs, such as pain (Spearman's correlation coefficient  $r_s=.62$ ), and weak correlations for SF-36 domains not included in the KOOS, such as mental health (Koos and Toksvig-Larsen, 2003). A further benefit of the KOOS is that the score houses the established Western Ontario and McMaster Universities Osteoarthritis (WOMAC) Index (Bellamy *et al.*, 1988) within it enabling WOMAC index scores to also be calculated.

*Reliability.* Test – retest data were tested in in one study of patients undergoing knee arthroplasty (Koos and Toksvig-Larsen, 2003). Intraclass correlation coefficients ( $ICC_{2,1}$ ) ranged from .78-.97 across the domains.

*Responsiveness to change.* Effects sizes across the subscales ranged from 1.24-2.86 at six months and 1.59-3.54 at one year after knee arthroplasty (Koos and Toksvig-Larsen, 2003); the most responsive scale being quality of life and the least responsive scale being the sport and recreation subscale.

*Floor and Ceiling effects.* In a study of 105 knee joint arthroplasty patients from pre operation up to one year post operation floor effects across the domains were generally between 0-16

percent, except for the sport and recreation domain which showed a 48% floor effect at pre operation (Koos and Toksvig-Larsen, 2003). Ceiling effects ranged from 1-22%.

#### **4.5.6.4 The Leg extensor press (LEP).**

*Purpose of the measure and reason for its inclusion in the trial.* The generation of adequate power, the rate at which muscles develop force, is required to successfully achieve functional tasks (Activity and Health Research Limited, 1990). Power is the product of force and velocity of contraction and the Leg extensor press which measures the explosive power in a single leg extension (Bassey et al., 1990; Frost et al., 2002) was included in this study. The intention of using the LEP was to provide a measurement relevant for functional activities which require speed as well as force (such as sit to stand) as well as slower controlled functional activities (such as stair climbing) (Bassey and Short, 1990).

*Measurement Approach and Scoring.* The LEP is designed to be a safe method of measuring explosive power, unlike other methods such as two-legged jumping on a force plate (Bassey and Short, 1990). The seat position is adjusted for leg length (Bassey and Short, 1990); important both for the production of power and to allow for the presence of fixed flexion deformities of the knee joint in our trial participants. Adherence to a strict test protocol to minimise measurement errors is required (Robertson et al., 1998).

The mean number of repetitions required for participants to reach their maximum LEP has been reported as 5.2 (Robertson et al., 1998; Bassey and Short, 1990). Usually up to maximum of 10 repetitions, with 15 second rest periods, are used for patients with knee joint pathology (Robertson et al., 1998; Frost et al., 2002). Since a relationship has been found between pain and function (Lamb et al., 1995; Robertson et al., 1998) it was predicted that

not all participants would be able to achieve 10 repetitions. The most common scoring approach is to use the highest recorded power value for each participant (Robertson *et al.*, 1998; Bassey *et al.*, 1992; Frost *et al.*, 2002).

*Validity.* LEP measurements in the elderly have previously found to be strongly associated with functional activities such as stair climbing (Bassey *et al.*, 1992). The LEP has also previously been used in a physiotherapy trial comparing exercise programmes following knee arthroplasty (Frost *et al.*, 2002) which, given the lack of comparable trials available, is considered advantageous.

*Reliability.* The level of repeatability of using the LEP for patients with osteoarthritis of the knee has been shown to be acceptable with coefficients of variation ranging from 9-16% being reported for this patient group (Robertson *et al.*, 1998). The same researchers also report Bland and Altman method for repeatability results; error of the standard deviation = .17 (95% of differences expected within  $\pm .47$ ).

#### **4.5.6.5 The Timed Walk Test.**

*Purpose of the Score and reason for its inclusion in the trial.* Timed walk tests are common outcomes in randomised clinical trials (Steadman *et al.*, 2003); measuring functional physical activity in the elderly for conditions such as osteoarthritis (Lin *et al.*, 2001) and, like this trial, to measure change physical function following knee arthroplasty (Shields *et al.*, 1995). Unlike the LEP, the timed walk and the STS below are both physical function tests which incorporate a number of elements such as strength, power and stability.

*Scoring Approach.* The test can be either the distance walked during a set time (such as 6 or 12 minutes) or the time taken to walk a set distance. In this trial the latter approach was chosen since patients may be unable to walk for long durations pre operatively or soon after surgery. The set distance used in previous research has varied from 2.44 m (Shaubert and Bohannon, 2005) to 8 feet (Lin *et al.*, 2001), 3 or 4 m (Wolinsky *et al.*, 2005) to 4.57 m (Shields *et al.*, 1995). A commonly used distance is 10 metres (Steadman *et al.*, 2003; Whitehead *et al.*, 2003). This distance was chosen since it was believed to be achievable for trial participants whilst being of sufficient duration to minimise the impact of any small stopwatch start/stop measurement errors.

Timed walk tests can be performed as fast as possible (Lin *et al.*, 2001) or at a normal comfortable speed (Shaubert and Bohannon, 2005; Wolinsky *et al.*, 2005). To optimise patient safety in the immediate post operative period, the latter approach was chosen. No verbal encouragement was used since this is known to influence test performance (Guyatt *et al.*, 1984).

*Validity.* Walking speed has long been acknowledged and accepted as a valid outcome measure following lower limb joint surgery; it is known to be a key measure for the analysis of gait (Bowker and Messenger, 1988).

*Reliability.* The coefficient of repeatability for an 8 feet test at natural speed been reported as .44 seconds with an intraclass correlation coefficient of 0.96 (Lin *et al.*, 2001). A variation of 7.3% has been found over nine maximum speed 10m tests which the authors considered to be relatively inconsequential when compared to normal biological systems (Howe and Oldham, 1995).



#### **4.5.6.6 The timed sit to stand (STS) test.**

*Purpose of the Measure and reason for its inclusion in the trial.* The sit-to-stand (STS) test is an objective functional measure of strength (Schaubert and Bohannon, 2005). It is included in the trial both as an objective measure of function, since it is fundamental to transfers and the initiation of gait (Eriksrud and Bohannon, 2003), and because of its correlation with other functional activities (see below).

*Measurement approach and scoring.* STS tests either measure the time taken to complete a specified number of STSs (Lin *et al.*, 2001; Wolinsky *et al.*, 2005) or record the number of STSs within a specified time period (Jones *et al.*, 1999). The former can be problematic in that a floor effect is often present with participants not being able to perform the 5 or 10 required number of STSs (Jones *et al.*, 1999). Previous research assessing physical function in the elderly found 24% (n=25) of participants with lower limb osteoarthritis were unable to complete this form of the test (Lin *et al.*, 2001).

Chair height influences the performance of the movement (Janssen *et al.*, 2002) so the same chair was used throughout the trial; used with its back legs braced against a wall to minimise chair movement (Jones *et al.*, 1999). Since patients with poor quadriceps strength are believed to compensate for this weakness by using their arms to get in and out of a chair (Lin *et al.*, 2001) patients were asked to perform the test with their arms folded. Although arm folding may, by influencing the body's centre of mass, require more adjustment at the ankle joint by patients than their usual movement (Carr, 1992) this approach was used because it has been calculated that the reduction of the extension moment required to STS is up to 50% if arm rests are used (Janssen *et al.*, 2002).

*Validity.* The STS has been moderately correlated ( $r=0.71-0.78$ ) to weight adjusted 1-RM (maximum) leg-press strength (Jones *et al.*, 1999) although it is appreciated that other factors, such as sensation, speed, balance and psychological status, are involved in its execution. (Lord *et al.*, 2002). The STS also correlates with knee extensor strength, ambulatory independence, walking speed and stair climbing performance (Bohannon, 1998; Bohannon and Eriksrud, 2001; Ikezoe *et al.*, 1997) and is a predictor of hospital length of stay post arthroplasty (Barker *et al.*, 2003). It is known that knee arthroplasty patients find STS a difficult manoeuvre compared to normals of the same age group and compensate, both pre and post operation, by taking more weight through the non affected side and by increasing forward leaning (Su *et al.*, 1998).

*Reliability.* The test-retest reliability for the number of STSs performed in 30s has been reported to be high ( $r=0.89$ ) and enables wide variations in ability levels to be reliably assessed (Jones *et al.*, 1999). The coefficient of repeatability for 5 STSs performed as fast as possible has been reported as 5.74 seconds, with an intraclass correlation coefficient of 0.96 (Lin *et al.*, 2001).

#### **4.5.6.7 Additional Measures.**

In addition to outcomes the following tools were included to inform the interpretation of the trial results and assist the assessment of internal validity of the trial.

#### ***The Intervention.***

The physiotherapists providing the intervention completed a case report form for each home visit immediately following each treatment (see Appendix XIII). The form recorded the time spent on the visit, and asked the physiotherapist to estimate the time (in minutes) spent upon each element of the intervention. Space was available for physiotherapists to add in any

additional treatment components incorporated into a treatment and to record any concerns participants raised. I have developed and successfully used this approach previously to check adherence to a pragmatic physiotherapy trial protocol (Mullis *et al.* 2006).

### ***Patient diaries.***

A patient resource use diary (Rivero-Arias *et al.*, 2005) was used to enable participants to record any additional non trial physiotherapy received following discharge from hospital and 12 month follow up (Appendix XII). Location, number of sessions, and source (NHS or privately funded) were recorded.

### ***Joint range of motion.***

The value of using joint range of knee range of motion as an outcome measure in rehabilitation trials is questionable (Minns Lowe *et al.*, 2007). Many variables influence post operative range of motion in addition to rehabilitation and these limit its usefulness as a measure of rehabilitation (Minns Lowe *et al.*, 2007): Lam *et al.*, (2003) summarise these variables and list implant design, surgical technique and approach and previous level of fixed flexion deformity in addition to rehabilitation. Joint range of motion therefore was not included as an outcome in this trial. However joint range has been recorded in all previous physiotherapy knee arthroplasty rehabilitation trials (Codine *et al.*, 2004; Frost *et al.*, 2002; Kramer *et al.*, 2003; Mockford and Beverland, 2004; Rajan *et al.*, 2004) bar one (Moffet *et al.*, 2004). To facilitate comparison with these other trials and because knee range of motion does affects performance and functional ability (Rowe *et al.*, 2000) and is a useful characteristic to meaningfully describe trial participants as required (Mant, 1999), this measurement was also included in the trial. Knee joint range of motion was measured via goniometry, a quick, easy, inexpensive way to measure motion which is more reliable than methods such as visual

estimation (Watkins *et al.*, 1991; Brosseau *et al.*, 2001) and which was also frequently used in the other trials.

#### **4.5.7 PROCEDURE AND INTERVENTION:**

##### **4.5.7.1 STAGE I. Recruitment of patients.**

An invite letter and trial information sheet were posted to patients prior to their attendance at the Pre Operative Admission Clinic (POAC). This was usually 10-14 days prior to attendance. Patients were requested to either return a reply slip in a pre-paid envelope if they were interested in participating or to inform the POAC staff at their appointment. The trial was fully verbally explained to interested patients at the POAC and informed valid consent sought.

##### *Baseline Assessment Procedures.*

Baseline data was collected for participants using self report questionnaires and standardised data recording forms (See Appendix XII). If, for any reason, only partial baseline measurements were obtained the reasons were recorded. Participants were also provided with their resource use patient diaries.

##### *Self Report Questionnaires.*

Patients were requested to read the questionnaires and to follow the instructions outlined. Participants were provided with a black pen to complete the questionnaire since this colour pen is read most accurately by the Optical Mark Recognition software (Remark Office version 6) used in the trial.

### *Joint Range of motion.*

Knee joint range of motion was measured using the approach of Norkin and White (2003). The standardised explanation, instructions to participants, and procedure can be found in Appendix XIV.

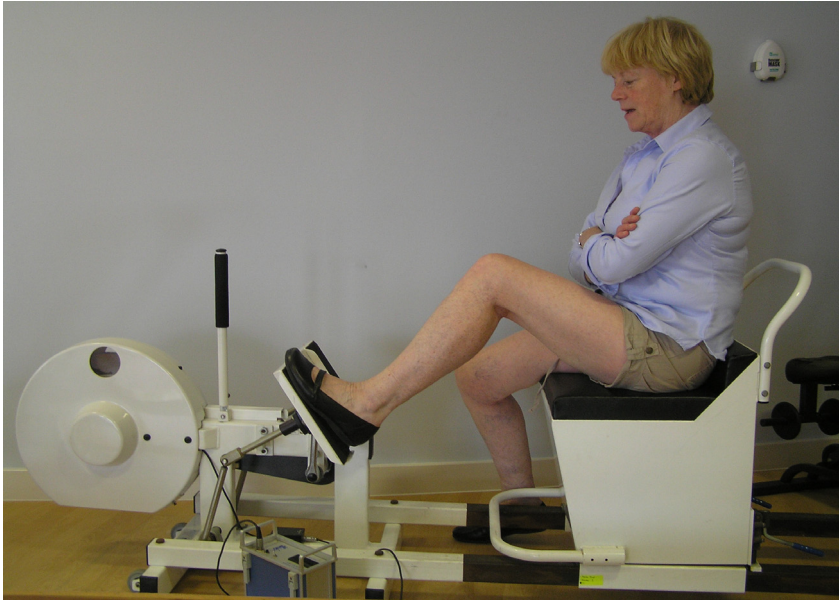
### *Timed Walk Test.*

A pre-measured 10m walkway was used for the test. During the test the assessor stood to one side and slightly behind the patient (to avoid pacing). A count down of “three, two, one, go.” was used. Timing commenced upon “go” and ceased when the patient crossed the 10m mark. The time was recorded in seconds (to two hundredths of a second) in the patient assessment questionnaire. All participants wore shoes for the test and were allowed to use their usual walking aid. If used, the type of aid was recorded in the patient assessment questionnaire. The standardised explanation, instructions to participants, and full procedure can be found in Appendix XIV.

### *Leg Extensor Press.*

The patient sat on the LEP seat with their back was positioned against the back rest. All participants wore shoes and folded their arms during the tests. The seat position was adjusted for each participant by their placing one foot on the fully depressed footplate and pushing the seat back slowly until the leg was extended as fully as comfortable possible. The seat position was clamped in place after adjustment and measured and recorded in centimetres (up to 2 decimal places) on the questionnaire. Measurement was from the distance from the front base

of the seat (labelled yellow marker 1) to the back end of the runner (labelled yellow marker 2) and this measurement was used throughout the baseline plus all follow up assessments.



**Illustration 19. Photograph showing the leg extensor press in use.**

Measurements were obtained for both legs, starting with the leg not undergoing knee replacement. Participants were instructed to place their foot upon the footplate with their free foot resting on the floor and push the foot pedal as hard and as fast as possible after the count down of “three, two, one, push”. Two warm up attempts were allowed for each leg. A relaxation period of 15 seconds occurred between each test. If pain did not halt testing, at least five tests (not including the warm ups) were carried out on each leg and recorded. If the patient still showed signs of improvement they were allowed to continue up to a maximum of 10 tests per leg if they were able to do so. The weight of each person on calibrated SECA bathroom scales was obtained and recorded in Kgs to use in analysis. The standardised explanation, instructions to participants, and full procedure can be found in Appendix XIV.

#### *Timed Sit-To-Stand.*

The standardised chair seat height for this test was 44 cm (including 4 cm of foam padding) and the seat depth was also 44 cm. The chair was without armrests. All participants wore shoes for this test. Participants started the test in the sitting position and were asked to perform the test with their arms folded. A count down of “three, two, one, go” was used and timing commenced upon the word go. No verbal encouragement was given. The number of times the patient completed a sit to stand was counted until 30 seconds passed at which time the timing was stopped and the patient told to cease the test. The number of completed sit-to-stands was recorded; no rating of partial movements was included. If the patient was unable to do any sit-to-stands then 0 was recorded.

The standardised explanation, instructions to participants, and full procedure can be found in Appendix XIV.

#### **4.5.7.2 STAGE II. Randomisation and Allocation Concealment.**

The well established randomisation procedure outlined by Pocock (1983 p. 66-79) was used. A randomisation list was prepared using random permuted blocks to prevent wide imbalances between the number of participants allocated to each treatment arm. A block size of 10 was used since this was considered sufficiently large to reduce predictability whilst ensuring that the number of participants requiring the intervention at any one time remained manageable. Randomisation was independently performed by the Director of the Physiotherapy Research Unit and the physiotherapists assessing and treating patients were not involved. The randomisation list was transferred to a sequence of sealed envelopes to promote allocation concealment from participants and assessing physiotherapists during baseline

assessments prior to randomisation, thus minimising selection bias (Altman *et al.*, 2001). Each envelope was sequentially numbered and contained the name of the next treatment on a card. Envelopes were opaque and were sellotaped after sealing. The envelopes were securely stored in a locked office and used in strict consecutive order. Upon completion of baseline collection measures, the next envelope was fetched and opened by the patient and replaced in the envelope by them. The study number of the patient was then written on the resealed envelope and returned to the trial co-ordinator.

#### **4.5.7.3 STAGE III. Surgery.**

Patients received elective surgery and usual hospital care via the implementation of the integrated care pathway already in use at the Nuffield Orthopaedic Centre, Oxford. Each patient's operation notes were reviewed and the grade of surgeon performing the operation noted. Any operative or post operative complications were recorded from the operation notes and, at a later stage, during the follow up assessments.

#### **4.5.7.4 STAGE IV. The Intervention.**

The intervention group patients received 2 physiotherapy home visits with the intervention package earlier described. The date, time and duration of each visit was recorded in minutes. Following each visit the physiotherapist wrote up the visit and also estimated the time (in minutes) spent on each of the activities on a form designed for the purpose (Appendix XIII). This was to enable both adherence to the trial protocol to be checked and to enable the intervention to be sufficiently described for future readers of the research.

The Usual Care Group received no additional home visits.



#### **4.5.7.5 STAGE V. Follow up.**

All trial patients underwent reassessment, using all outcome measures performed at baseline, at three and twelve months post operatively by a physiotherapist blind to trial arm allocation. The three month timepoint was chosen as an end of treatment time, allowing time for participants to practice exercises and tasks practiced during the second visit. The final twelve month timepoint was chosen since the most improvement following knee arthroplasty occurs within the first year, with the largest gains being in the first six months (Jones *et al.*, 2007). All patients were also followed up at six months post operatively by postal questionnaire containing the Oxford Knee score and the KOOS. The use of postal assessment in this way has previously been demonstrated to be a feasible method for rehabilitation trials (Parker and Dewey, 2000). To promote compliance with full completion of the patient diaries, each participant was reminded to complete their diaries verbally at follow up visits, by telephone when arranging visits or following up patients by telephone if they were unable to attend the hospital for follow up and in the letter sent to the patient with their 6 month follow up questionnaire.

#### **4.5.8 DATA ANALYSIS.**

##### **4.5.8.1 Missing or multiple data.**

*Oxford Knee Score.* One or two missing values were substituted with the average value for the score; if more than two items were missing the score was deemed invalid (Murray *et al.*, 2007). If two boxes were marked the worst outcome was chosen (Murray *et al.*, 2007).

*KOOS.* One or two missing values were substituted with the average value for that subscale; if more than two items were missing the subscale was deemed invalid (KOOS User's Guide, 2003). If two boxes were marked the worst outcome was chosen (KOOS User's Guide, 2003).

All other missing data was assigned the value 999. Data for tests which participants attempted but were unable to perform were given a value of 0.

#### **4.5.8.2 Data Analysis.**

An intention to treat approach was used for data analysis. In this type of analysis the data from each participant are analysed in accordance with the treatment arm to which they have been randomised (Sim and Wright, 2000 p.330).

The results first underwent descriptive analysis to describe the characteristics of the sample and to check the variables for any violations of the assumptions underlying the planned statistical tests (Pallant, 2007 p. 53). The exclude cases pairwise option was chosen; this includes a case in all analyses for which they have data and only excludes a case if they are missing the data required for a specific analysis. Kolmogorov-Smirnov statistics were calculated to further assess the normality of the distribution scores and assist confirmation of significant deviations from the norm (Field, 2000 p.47-48).

The baseline characteristics were examined to ascertain the extent to which randomisation was effective and how similar the groups appeared at baseline (Altman, 1991 p.38-39). Following this descriptive analyses were performed for both trial arms at all time points using numbers and percentages for categorical variables and means and standard deviations for continuous variables (Altman, 1991 p.38-39). Within group comparisons of primary and secondary outcomes were performed between each timepoint: paired t-tests were used for the outcomes comparing normally distributed data and Wilcoxon Signed Ranks test for outcomes comparing non normally distributed data (Campbell and Machin, 1993, 75-84). Outcome scores between each group were then compared between each timepoint. Independent samples t-tests were used for the outcomes comparing normally distributed data

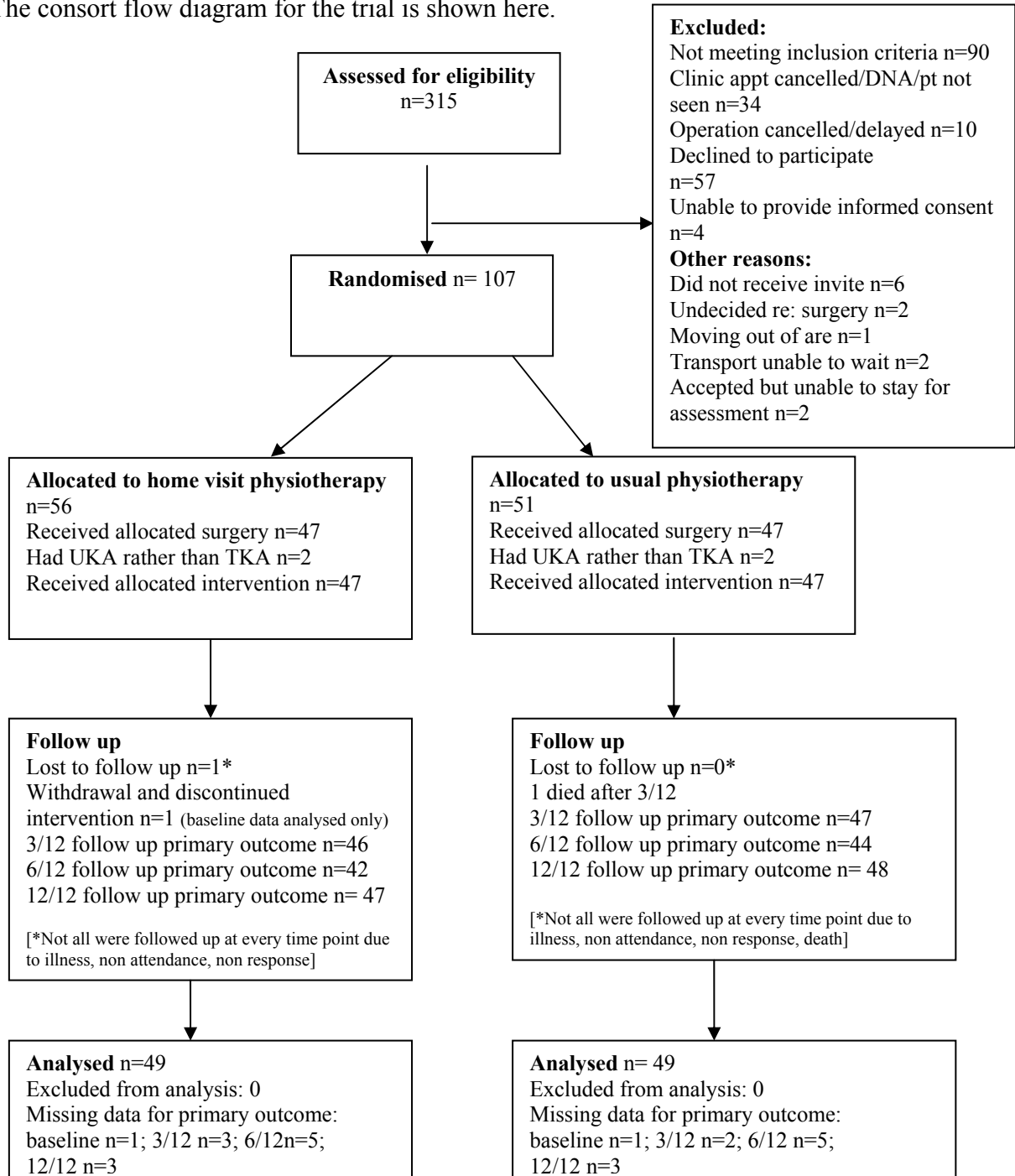
and Mann-Whitney U tests were deployed for outcomes comparing non normally distributed data (Altman, 1991 p.194-195). One-way repeated measures analysis of variance tests were then used. Since responses to treatment may show large variations by virtue of differences across individuals, this test allows responses to treatment to be compared for each person, thus minimising the variability across people (Sim and Wright, 2000 p. 309). These tests provide greater power to detect effects; however, since scores at each timepoint are likely to be related because they come from the same subjects, the accuracy of the F-test is adversely effected (Field, 2000 p. 323-326). This means that an additional assumption has to be made, of sphericity, where it is assumed that the relationship between pairs of experimental conditions is similar (Field, 2000 p. 323-326). To assist ease of readability and clarity further details regarding specific analyses are provided in the results section to prevent repeated flicking back and forwards over many pages.

In addition to making statistical comparisons of follow up scores, change scores were also calculated to allow the changes in outcome scores to be explicitly stated (Vickers and Altman, 2001). Finally, since subgroup analyses are unreliable in small studies, no subgroup analyses were undertaken in this trial (Friedman *et al.*, 1998 p. 304-306).

## 4.6 RESULTS.

### 4.6.1 FLOW OF PARTICIPANTS THROUGH THE TRIAL.

The consort flow diagram for the trial is shown here.



**Figure 13. CONSORT flow diagram,**

Although patients were not asked to provide a reason for declining to participate in the trial 42% of those declining spontaneously cited travel and transport issues as their prime reason for non participation because the follow up visits necessitated 2 extra hospital visits. A further two potential participants could not be seen in clinic due to hospital transport time constraints.

#### **4.6.2 COMORBIDITIES, MORTALITY AND COMPLICATION RATES REPORTED DURING THE TRIAL.**

##### **4.6.2.1 Comorbidity and mortality.**

In the home visit group one patient was diagnosed and treated for multiple myeloma, another was diagnosed and operated upon for hydrocephalus, one participant was diagnosed and treated for rheumatoid arthritis and a further participant had been referred to the spinal team for possible surgery for stenosis. One participant in the home treatment group became progressively cognitively impaired following surgery and unable to comply with full follow up assessment procedures.

Overall less co-morbidities were developed by usual physiotherapy group participants. One participant was diagnosed and treated for anaemia and one participant reported progressive sensory loss due to diabetes. However one patient in the usual physiotherapy group died during the trial; the cause of death was unrelated to their knee operation.

#### **4.6.2.2 Complications Reported By Participants At Three Month and Twelve Month Follow Up Assessments.**

Eleven participants reported a total number of 17 major complications occurring in the first three months. A total of 83 minor complications were reported by 56 participants with a minority of 36 participants reporting no minor complications. Details can be found in the tables below.

The reported number of major complications within the first three months was checked against the medical records for the relevant participants to enable comparison of rates with existing published data from available publications. The medical records made no mention of deep infections for the two participants reporting them, and no mention of the two DVTs. The records confirmed the presence of a pulmonary embolus for the participant mentioning this (and also interestingly records the presence of a DVT the participant had not mentioned). The reported intra-operative fracture was confirmed by the medical notes. The three chest infections reported by participants were not mentioned in their medical records. The myocardial infarction was confirmed as was the foot drop with associated nerve damage. Of the three strokes/trans ischaemic attacks reported, one major stroke was confirmed, one TIA was confirmed whilst the medical records for the third were not available. Of the three pressure sores reported by patients, whilst only one was confirmed by medical records, confirmation of the others was obtained via the patient resource diaries which listed multiple visits to have pressure sores dressed.

The number of major complications reported from the start of month four until the end of twelve months increased to 24 in 22 participants, largely due to 16 participants reporting chest infection during this time. Importantly two revisions had already taken place with a further two participants exploring the possibility of further surgery.

**Table 8. Post operative major complications related to knee replacement surgery for 0-3 months (n=92) and 4-12 months following surgery (n=93).**

Complication	0-3 Month Follow up		4-12 Month Follow up	
	No of participants with complications/ no of participants for which data was obtained	Percentage of no of participants for which data was obtained	No of participants with complications/ no of participants for which data was obtained	Percentage of no of participants for which data was obtained
Deep infection	2/91	2.2%	0	
Deep vein thrombosis	2/91	2.2%	0	
Pulmonary embolus (requiring anticoagulants)	1/92	1.1%	0	
Intra-operative fracture	1/92	1.1%	0	
Chest infection	3/92	3.3%	16/93	17.2%
Heart attack	1/92	1.1%	1/92	1.09%
Stroke /TIA	3/92	3.3%	2/92	2.17%
Pressure sores	3/92	3.3%	0	
Nerve damage with foot drop	1/92	1.1%	1/89	1.12%
Patellar button moved – awaiting further investigations			1/89	1.12%
TKA revisions			2/89 [1: tibial component slipped, revised, fractured 2/52 later 2: sudden onset burning pain- TKA required revision]	2.25%
Daily knee locking – awaiting investigation			1/89	1.12%
<i>Total number of major complications</i>	<i>17 in 11 participants</i>	<i>Overall major complication rate =18.48%</i> <i>Percentage of participants reporting major complications = 11.96%</i>	<i>24 in 22 participants</i>	<i>Overall major complication rate =25.81%</i> <i>Percentage of participants reporting major complications = 23.66%</i>

The reported number of major complications were then separated into home and usual physiotherapy groups and the results can be seen in the table below. Overall, the rates

between the two groups appeared similar although both revisions and the majority of strokes were in the home treatment group.

**Table 9. Table to show treatment arm allocation for post operative major complications related to knee replacement surgery for 0-3 months (n=92) and 4-12 months following surgery (n=93).**

Complication	0-3 Month Follow up		4-12 Month Follow up	
	No of participants in home visit group	No of participants in usual physiotherapy group	No of participants in home visit group	No of participants in usual physiotherapy group
Deep infection	1	1	0	
Deep vein thrombosis	1	1	0	
Pulmonary embolus (requiring anticoagulants)	1		0	
Intra-operative fracture	1	0	0	
Chest infection	1	2	8	8
Heart attack	0	1	0	1
Stroke /TIA	2	1	2	0
Pressure sores	1	2	0	
Nerve damage with foot drop	0	1	0	1
Patellar button moved – awaiting further investigations			1	
TKA revisions			2	0
Daily knee locking – awaiting investigation			0	1
<i>Total number of major complications</i>	<i>8</i>	<i>9</i>	<i>13</i>	<i>11</i>

The minor complications reported by participants can be seen in the table below. The number of minor complications reported lessened over time, with 32 complications reported by a minority of 23 patients. The most common of these were 7 participants complaining of continued knee pain and 6 with swelling/inflammation.



**Table 10. Post operative minor complications related to knee replacement surgery within three months following surgery according to patients (self report closed and open question data n=92).**

Complications	0-3 Months Follow Up Number of participants reporting complication (Percentage of number of participants for which data was obtained for closed questions only)	4-12 Months Follow Up Number of participants reporting complication (Percentage of number of participants for which data was obtained for closed questions only)
Superficial wound infection (requiring antibiotics)	15 (16.3%)	4 (4.3%)
Haemorrhage requiring additional/unexpected transfusion	10 (10.9%)	
Clips/Stitches (undissolved, left in, difficult to remove, causing ulcer)	6	
Excessive pain/nerve pain &/or bruising	6	
Falls	5	
Swelling & inflammation	4	6
Limited flexion/ROM/MUA	4	
Ankle pain and/or swelling limiting mobility	3	
Drain not working/falling out	3	
Haemarthrosis, haematoma	3	
Anaemia requiring medication (not transfusion)	2	
Blisters	2	
Cellulitis	2	
Depression, panic attacks	2	1
Heat in knee since operation	2	1
Numbness	2	1
Chest pain	1	
Angina	1	
Foot coldness since operation	1	
Hamstring clicking during extension	1	
Ineffectual nerve block	1	
Joint clicking with pain	1	
Leg length discrepancy	1	
Post operative hypotension	1	
Problematic scar tissue	1	
Stomach bug	1	
Varicose eczema	1	1
Wound Bleeding	1	
Continued pain		7
Knee cap rubbing on prosthesis		1
Tibia and foot pain – awaiting scan results		1
Knee itchiness		1
Knee ligament injury after twisting knee		1
Intermittent redness		1
Opposite knee arthritis		1
Partial Achilles rupture opposite leg		1
Knee giving way		1
Rash on knee		1
Ruptured Bakers cyst		1
Scar tissue in lateral ligament		1
<b>Total</b>	<b>83 complications in 56 participants</b>	<b>32 complications in 23 participants</b>

#### **4.6.3 Further Interventions and Hospital Admissions Following Surgery.**

The table below shows the number of further interventions or operations and further admissions required by participants during the trial. Six participants required further interventions in the first three months following their surgery. Three of these participants underwent knee joint manipulations under anaesthetic (home visit group n=1, usual physiotherapy group n=2), one home visit participant had a clip removed from the bottom of their wound (patient admitted three times into the bone infection unit), one usual physiotherapy underwent nerve conduction studies and one home visit had an MRI for severe back and leg pain following surgery. Seven further patients were admitted into hospital at least once during the first three months following surgery: one following discharge with a DVT and pulmonary embolus (home visit group), one after a fall (n=1 usual physiotherapy group), one after a slip and major stroke (n=1 home visit group), bleed into joint (n=1 usual care group), knee pain (n=1 home visit group), pulled tendon (n=1 home visit group) and chest pain (n=1, three admissions, home visit group).

Five participants required further interventions from the start of month 4 to twelve months post surgery; two patients in the home visit group underwent revisions, one in the usual physiotherapy group had a biopsy (results not known), one usual physiotherapy participant had fluid removed from the knee and one usual physiotherapy participant reported 3-4 courses of antibiotics.

Four patients were admitted into hospital between 4-12 twelve months, two for revision arthroplasties, one with multiple myeloma and one due to rheumatoid arthritis. All were in the home visit group.

**Table 11. Number of self reported interventions and hospital admissions within 0-3 (n=93) and 4-12 months (n=93) following surgery.**

Intervention/Admission	0-3 months	4-12 months
	Number of participants with complication/Number of participants for which data was obtained	Number of participants with complication/Number of participants for which data was obtained
Number of further operations or interventions related to this joint replacement only	6/93 or 6.45%	5/92
Post operative complication/s which required the patient to be transferred to another hospital	1/93 or 1.1%	0
Number of readmissions to hospital	12/93 (total of 17 admissions) or 12.9%	4/93 or 4.17%

#### **4.6.4 Further Joint Replacement Surgery.**

In addition to the two revisions already mentioned, five participants underwent total knee arthroplasties to their other knee (n=4 home visit and one usual care participant), and one further participant a unilateral knee joint arthroplasty to their other knee (usual physiotherapy group), during the twelve month follow up period. Three other participants reported hip procedures: two hip arthroplasties (both home visit group) and one hip resurfacing procedure (usual care group). Overall, of the 12 procedures, 9 were undertaken by home visit group participants and 3 by usual physiotherapy group participants.

#### **4.6.5 BASELINE CHARACTERISTICS.**

##### **4.6.5.1 Descriptive statistics for the primary outcome at baseline.**

Descriptive statistics were used to describe the characteristics of the sample and to check the variables for any violations of the assumptions underlying the planned statistical tests (Pallant, 2007 p. 53). The exclude cases pairwise option was chosen; this includes a case in all analyses

for which they have data and only excludes a case if they are missing the data required for a specific analysis. Appendix XV contains additional information regarding all the results provided in this chapter.

***Initial description of the baseline primary outcome data.***

The table below provides an initial description of the baseline data for the primary outcome, the Oxford Knee Score. Baseline scores show wide variation from 4-41 out of a possible 0-48; plus the kurtosis value of -0.245 also indicates a slightly flattened distribution (Pallant, 2007 p. 56). The 5% trimmed mean value of 19.98 is close to the mean of 20.11, indicating that extreme scores are unlikely to be unduly influencing the mean (Pallant, 2007 p. 59). A table of extreme values is in Appendix XV. The standard deviation of 7.857 is lower than other published data for this patient group (Dawson *et al.*, 1998). The slight positive skewness value of 0.145 indicates an absence of serious skew problems with this data (zero indicates normally distributed data) (Altman, 1991 p. 36).

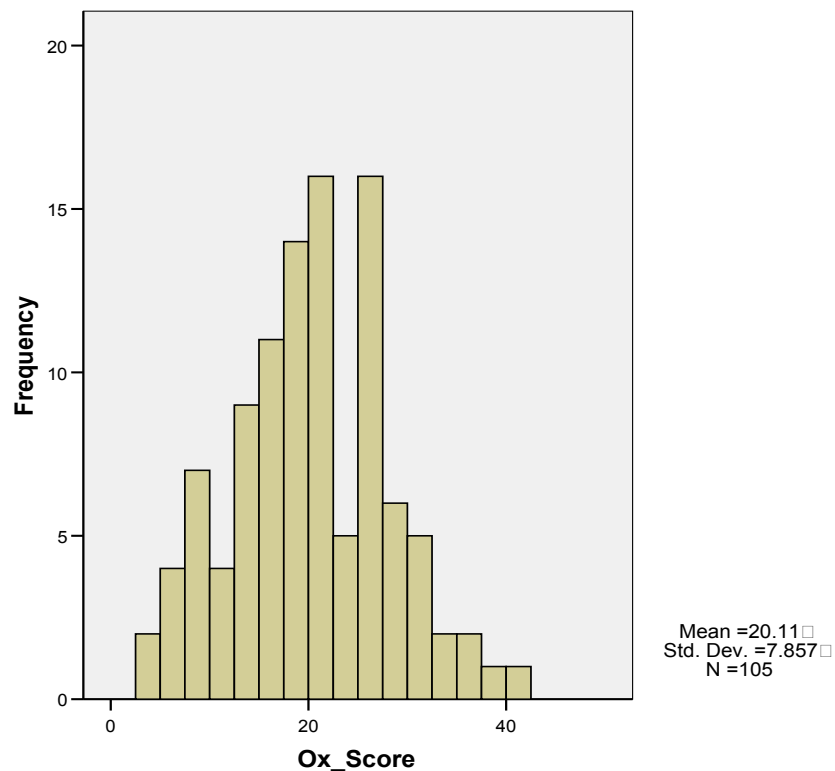
**Table 10. Table to describe the data obtained for the primary outcome at baseline (n=105).**

			Statistic	Std. Error
Ox_Score	Mean		20.11	.767
	95% Confidence Interval for Mean	Lower Bound	18.59	
		Upper Bound	21.63	
	5% Trimmed Mean		19.98	
	Median		20.00	
	Std. Deviation		7.857	
	Minimum		4	
	Maximum		41	
	Interquartile Range		11	
	Skewness		.145	.236
	Kurtosis		-.245	.467

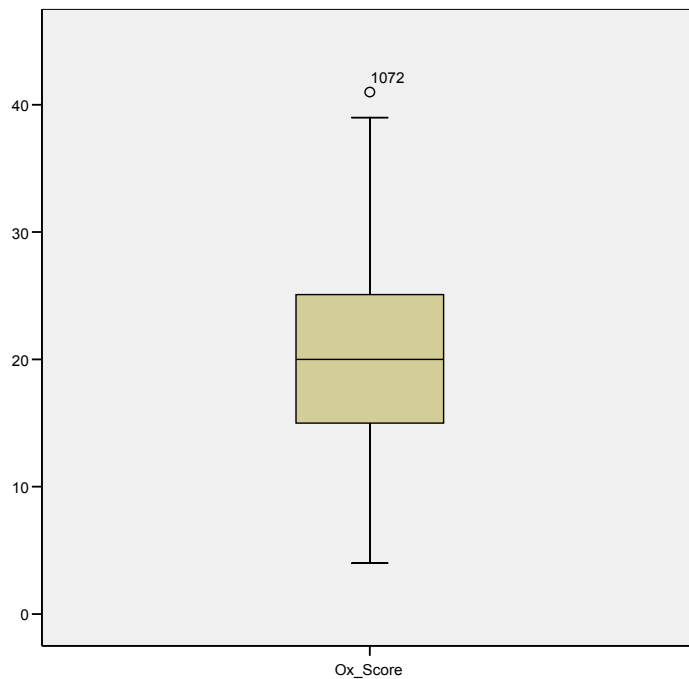
***Further assessment of Normality within the baseline Oxford Knee Score data.***

The histogram below shows the distribution of scores for the primary outcome, the Oxford Knee Score. Apart from the dip observed around scores of 24-26, the data on the histogram appears approximately normally distributed. A Kolmogorov-Smirnov statistic was calculated to further assess the normality of the distribution scores (Appendix XV). The statistic value of 0.048 with 105 degrees of freedom gave a lower bound value for the true significance of 0.200. This non – significant value indicates data normality.

The box-whisker plot below does identify one outlier (ID 1072). Since this outlier is close to the upper whisker it is not considered extreme and the decision was made not to exclude this point from analysis.



**Figure 14. Histogram to show the distribution of scores at baseline for the primary outcome, the Oxford Knee Score (n=105).**



**Figure 15. Box-whisker plot of the Oxford Knee Score data showing the 2 ½ , 25, 50, 75 and 95 ½ cumulative relative frequencies (centiles) plus one outlier ID 1072.**

*Assessment of Normality for the secondary outcomes at baseline.*

The distribution of data was also explored for the secondary outcomes (see Appendix XV). The symptoms, pain, functional activities of daily living, subscales of the KOOS provided data which were normally distributed. The sport and recreation and quality of life subscales data were not normally distributed. Similarly the data from the maximum leg extensor press values/weight in kg, for the timed walk test and for the timed sit to stand were also not normally distributed.

#### **4.6.6 Baseline characteristics for the two trial groups.**

The baseline characteristics for the two groups of the trial are summarised in the table below. The two trial arms generally appear relatively similar in terms of baseline characteristics although more participants in the home visit group had their left knee replaced whilst more participants in the usual physiotherapy group had their right knee replaced. Although it might appear that the mean duration of symptoms is greater for the usual physiotherapy group than for the home visit group, when months are converted to years the groups demonstrate greater similarity (home visit group mean = 11.2 years; usual care group mean = 12.3 years). Since there were some extreme outliers present the median value is also provided for symptom duration. The p values are included solely as a means of checking that effective randomisation occurred in the trial. Since none of the p values show significant differences between the groups it is assumed that randomisation was successful.

**Table 11. Baseline characteristics of the Home Visit and the Usual Physiotherapy Groups.**

Characteristics at Baseline	Home Visit Group (no of participants providing data)	Usual Physiotherapy (no of participants providing data)	P values t-test for equality of means <sup>#</sup> OR Mann Witney U for non parametric <sup>§</sup> OR Chi-square*
Total number of participants randomised to each group	56	51	
Gender: males females	24 32	21 30	1.00*
Mean age (years) Standard Deviation	67.84 (n=56) SD (8.45)	70.76 (n=51) SD (9.45)	.094 <sup>#</sup>
Knee being replaced: left right undecided (no operation)	31 24 1	22 29	**
Median Functional Comorbidity Index Interquartile range	2 IQ =1	2 IQ=1	.611 <sup>§</sup>
Mean Body Mass Index Standard deviation	31.32 (n=56) SD 6.28	29.27 (n=51) SD 5.82	.082 <sup>#</sup>
Mean Duration of symptoms (months) Standard deviation Median Duration of symptoms (months)	133.91 (n=55) 141.53 84	147.92 (n=51) 146.84 96	.618 <sup>#</sup>
Mean Range of max knee joint flexion (degrees) Standard deviation	105.26 (n=53) SD 10.15	105.38 (n=51) SD 19.92	.976 <sup>#</sup>
Mean Range minimum flexion (degrees) Standard deviation	9.27 (n=53) SD 11.464	9.93 (n=51) SD 8.732	.810 <sup>§</sup>
Number of participants with symptoms in the other knee	37 (n=53)	29 (n=51)	.243*
Number of participants with previous other arthroplasties	21 (n=55)	22 (n=51)	.419*
Number of participants with a history of diabetes	7 (n=55)	3 (n=51)	**
Number of participants with a history of cardiac disease	13 (n=55)	17 (n=50)	.338*
Number of participants who have ever smoked	28 (n=55)	25 (n=50)	1.00
Number of participants who currently smoke	4 (n=56)	2 (n=51)	**
<i>Number of participants who currently drink alcohol</i>	39 (n=53)	34 (n=50)	.684

\*\* Since 1 or more cells had an expected count less than 5 these Chi-square tests were violated and values are not provided.



#### **4.6.7 Baseline Outcomes for each group.**

The baseline outcomes for the two groups of the trial are summarised in the table below. The two trial arms appeared relatively similar in terms of baseline outcomes. Once again, the p values are included solely as a means of checking that effective randomisation occurred in the trial. The tests were not performed to see if the groups were random samples from the same population (the null hypothesis) since this was already known. Also, because the groups were allocated at random, it was already known that any differences are due to chance variation alone. Since none of the p values show significant differences between the groups it is assumed that randomisation was successful.

**Table 12. Baseline Outcome scores for the Home Visit and the Usual Physiotherapy Groups.**

Baseline Outcome scores	Home Visit Group (no of participants providing data)	Usual Physiotherapy (no of participants providing data)	P values t-test for equality of means <sup>#</sup> OR Mann Witney U for non parametric <sup>§</sup>
Mean Oxford Knee Score Standard deviation	20.43 (n=55) SD 7.20	19.76 (n=50) SD 8.58	p=.66 <sup>#</sup>
Mean KOOS symptoms subscale Standard deviation	41.66 (n=56) SD 18.43	41.29 (n=51) SD 22.29	p=.36 <sup>#</sup>
Mean KOOS pain subscale Standard deviation	39.65 (n=56) SD 16.81	38.53 (n=51) SD 17.29	p=.74 <sup>#</sup>
Mean KOOS ADL subscale Standard Deviation	45.51 (n=54) SD 14.40	45.01 (n=48) SD 20.57	p=.89 <sup>#</sup>
Median KOOS sport & recreation subscale Interquartile range	12.5 (n=42) 25	10 (n=33) 37	p=.78 <sup>§</sup>
Median KOOS quality of life subscale Interquartile range	21.88 (n =56) 25	25.0 (n=50) 31	p=.99 <sup>§</sup>
Median leg extensor press maximum power values /weight in kg for operated leg Interquartile range	.39 (n=49) .48	.34 (n=48) .57	p=.882 <sup>§</sup>
Median leg extensor press maximum power values /weight in kg for non operated leg Interquartile range	.52 (n=51) .66	.57 (n=50) .67	p=.449 <sup>§</sup>
Median timed walk test Interquartile range	11.73 (n=54) 5.42	11.88 (n=51) 6.04	p=.974 <sup>§</sup>
Median no of completed STSs in 30 sec Interquartile range	4.5 (n=55) 5.42	4 (n=50) 6.04	p=.846 <sup>§</sup>

NB: OKS scoring 0-48 with 48 being the best outcome.

KOOS each domain scored 0-100 with 100 being the best outcome

## **4.7 POSTOPERATIVE FOLLOW UP DATA AT THREE, SIX AND 12 MONTH TIMEPOINTS.**

### **4.7.1 Assessment of Normality within the Oxford Knee Score primary outcome follow up data.**

As for the baseline data, the Oxford Knee Score follow up data underwent descriptive analysis (Appendix XV). The table below summarises these findings. The median values show steady improvement from baseline to 12 months, although the minimum values show that not everyone follows this trend. The next table presents the results from Kolmogorov-Smirnov statistical tests at each time point. Whilst a non significant value, indicating normality, was obtained for the baseline data the significant values for all subsequent time points indicate the data was not normally distributed throughout the follow up period.

**Table 13. Table to describe the data obtained for the primary outcome at all timepoints (n=105).**

	Baseline OKS (n=105)	3 months OKS (n=93)	6 months OKS (n=86)	12 months OKS (n=94)
Mean (standard error)	20.11 (.767)	32.48 (.898)	35.06 (.982)	36.82 (.919)
95% confidence interval for mean	18.59 to 21.63	30.69 to 34.26	33.11 to 37.01	34.99 to 38.64
5% trimmed mean	19.98	32.85	35.55	37.45
Median	20	34	36	40
Standard deviation	7.857	8.656	9.106	8.908
Minimum	4	9	11	8
Maximum	41	47	48	48
Interquartile range	11	13	12	10
Skewness (standard error)	.145 (.236)	-.628 (.25)	-.693 (.26)	-1.093 (.249)
Kurtosis (standard error)	-.245 (.467)	-.114 (.495)	-.011 (.514)	.674 (.493)

NB: OKS scoring 0-48 with 48 being the best outcome.

**Table 14. Kolmogorov-Smirnov statistical test results for the Oxford Knee Score at all timepoints.**

	Kolmogorov-Smirnov(a) statistic	Degrees of freedom	Significance
<b>Baseline</b>	.048	105	.200(*)
<b>3 months</b>	.099	93	.025
<b>6 months</b>	.101	86	.030
<b>12 months</b>	.150	94	<.000

\* This is a lower bound of the true significance.

#### **4.7.2 Distribution of secondary outcome follow up data.**

The distribution of follow up data was also explored for the secondary outcomes (see Appendix XV). No outcome was normally distributed across all timepoints. The KOOS symptoms subscale showed normal distributions at baseline and three months but not at 6 or 12 months. The KOOS pain and quality of life subscales showed normal distribution at 6 months but not at three or twelve months. The KOOS ADL subscale showed normal distributions at all timepoints except 12 months. The KOOS sports scale data showed normal distributions at 3, 6 and 12 months but not at baseline. The leg extensor press values showed normal distribution for 3/12 for not for 12/12. The other outcomes at all other timepoints were non normally distributed.

#### **4.7.3 Descriptive analysis of primary and secondary outcome follow up data.**

The table below summarises the descriptive analyses for outcome data for both trial arms at all time points; since the data were not normally distributed across all timepoints the medians and interquartile ranges are presented in this table (means are included in Appendix XV). In addition, for the few timepoints and outcomes where data was normally distributed the means and standard deviations have also been presented.

**Table 15. Outcome scores for the Home Visit (HV) and the Usual Physiotherapy Care (UPC) Groups at baseline, 3, 6, and 12 months.**

Outcome	Baseline		3months		6 months		12 months	
	H V (no of pts)	U PC	H V (no of pts)	U PC	HV (no of pts)	U PC	HV (no of pts)	U PC
Median Oxford Knee Score	20 (n=55)	19.5 (n=50)	33.5 (n=46)	34 (n=47)	36 (n=42)	36 (n=44)	40 (n=46)	38.5 (n=48)
IQ range	10	12	13	12	12	13	10	13
Mean KOOS symptoms subscale	41.66 (n=56)	41.29 (n=51)	66.07 (n=46)	68.34 (n=47)	(n=42)	(n=44)	(n=44)	(n=48)
Standard Deviation	18.43	22.29	17.90	19.01				
Median	39.29	39.29	67.86	71.43	76.79	71.43	82.14	78.79
IQ range	21	36	29	29	21	29	18	31
Median KOOS pain subscale	41.15 (n=56)	38.89 (n=51)	69.1 (n=46)	72.22 (n=47)	75 (n=42)	75 (n=43)	80.56 (n=44)	90.33 (n=48)
IQ	22	19	28	29	25	31	36	33
Mean KOOS ADL subscale	45.51 (n=54)	45.01 (n=48)	70.56 (n=42)	73.43 (n=44)				
Standard deviation	14.398	20.572	16.747	18.897	(n=41)	(n=41)	(n=41)	(n=46)
Median	45.45	43.93	69.85	75	78.13	72.06	85.29	89.43
IQ range	21	28	21	31	26	34	21	23
Median KOOS sport & recreation subscale, IQ	12.5 (n=42)	10 (n=33)	41.67 (n=18)	31.67 (n=22)	50 (n=28)	35 (n=30)	60 (n=33)	50 (n=25)
	25	37	25	57	44	35	41	55
Median KOOS quality of life subscale, IQ	21.88 (n=56)	25 (n=50)	53.13 (n=46)	56.25 (n=47)	59.38 (n=42)	59.38 (n=44)	63 (n=44)	62.5 (n=48)
	25	31	25	31	31	41	43	42
Median leg extensor press of maximum W/kg for operated leg IQ	.39 (n=49)	.335 (n=48)	.7 (n=42)	.72 (n=39)	N/A	N/A	.87 (n=38)	.865 (n=42)
	.48	.57	.85	.67			.77	1.07
Median leg extensor press of maximum W/kg for non operated leg IQ	.52 (n=51)	.57 (n=50)	.725 (n=42)	.88 (n=40)	N/A	N/A	.855 (n=36)	.83 (n=43)
	.66	.67	.9	.87			.76	1.03
Median timed walk test in seconds	11.72 (n=54)	11.88 (n=51)	9.93 (n=42)	10.26 (n=43)	N/A	N/A	9.18 (n=40)	9.09 (n=43)
IQ	5.42	6.04	2.74	4.18			3.01	2.37
Median number of completed sit to stands in 30 seconds	4.5 (n=55)	4 (n=50)	7 (n=43)	7 (n=43)	N/A	N/A	7 (n=41)	8 (n=43)
IQ	5	6.04	4	6			8	6

NB: OKS scoring 0-48 with 48 being the best outcome.

KOOS each domain scored 0-100 with 100 being the best outcome

#### **4.7.4 Within group comparisons of primary and secondary outcome follow up data.**

Outcome scores within each group were compared between each timepoint. Paired t-tests were used for the outcomes comparing normally distributed data and Wilcoxon Signed Ranks test for outcomes comparing non normally distributed data. The results can be seen in the table below.

This table shows that the within group change scores are significantly different at baseline and three months, and baseline and twelve months, in both groups for all outcomes. The values for both groups look similar.

At the 3-6 months time points the groups show some differences. The KOOS symptom subscale showed a significant value for the home visit group but not for the usual care group. The KOOS ADL and sport and recreation subscales values for each group are dissimilar. The Oxford knee score, KOOS pain and KOOS quality of life scores showed non significant changes.

At the 6-12 months timepoints there is a significant difference in Oxford Knee Scores for the home visit group but not for the usual physiotherapy group. Both groups show a significant difference in scores for the KOOS symptoms. The usual physiotherapy group shows significant differences in KOOS pain, functional activities of daily living and sports and recreational scores where as the home visit group does not. Neither group shows a significant difference for KOOS quality of life scores.

At the 3-12 months timepoints, for outcomes not collected at 6 months, both groups show significant differences for the leg extensor press scores and the timed walk scores and both show non significant differences for the time sit to stand scores.

**Table 16. Table presenting the comparisons of scores within groups at baseline, 3, 6, and 12/12 timepoints.**

Outcome	Baseline to 3/12 P values		3/12 to 6/12 P values		6/12 to 12/12 P values		Baseline to 12/12 P values	
	Paired t-test # OR Wilcoxon Signed Rank <sup>§</sup>		Paired t-test # OR Wilcoxon Signed Rank <sup>§</sup>		<b><i>OR 3 to 12/12 p values (in italics)</i></b> Paired t-test # OR Wilcoxon Signed Rank <sup>§</sup>		Wilcoxon Signed Rank <sup>§</sup>	
	Home Visit	Usual Physiotherapy	Home Visit	Usual Physiotherapy	Home Visit	Usual Physiotherapy	Home Visit	Usual Physiotherapy
Oxford Knee Score	p <.000 <sup>§</sup>	p <.000 <sup>§</sup>	.074 <sup>§</sup>	.095 <sup>§</sup>	.026 <sup>§</sup>	.080 <sup>§</sup>	p <.000 <sup>§</sup>	p <.000 <sup>§</sup>
KOOS symptoms	p <.000 <sup>#</sup>	p <.000 <sup>#</sup>	.038 <sup>§</sup>	.457 <sup>§</sup>	.007 <sup>§</sup>	.003 <sup>§</sup>	p <.000 <sup>§</sup>	p <.000 <sup>§</sup>
KOOS pain	p <.000 <sup>§</sup>	p <.000 <sup>§</sup>	.158 <sup>§</sup>	.109 <sup>§</sup>	.181 <sup>§</sup>	.002 <sup>§</sup>	p <.000 <sup>§</sup>	p <.000 <sup>§</sup>
KOOS ADL	p <.000 <sup>#</sup>	p <.000 <sup>#</sup>	.101 <sup>#</sup>	.871 <sup>#</sup>	.074 <sup>§</sup>	.003 <sup>§</sup>	p <.000 <sup>§</sup>	p <.000 <sup>§</sup>
KOOS sport & recreation	p <.000 <sup>§</sup>	p <.000 <sup>§</sup>	.099 <sup>#</sup>	.828 <sup>#</sup>	.123 <sup>#</sup>	.018 <sup>#</sup>	p <.000 <sup>§</sup>	p <.000 <sup>§</sup>
KOOS QoL	p <.000 <sup>§</sup>	p <.000 <sup>§</sup>	.471 <sup>§</sup>	.605 <sup>§</sup>	.080 <sup>§</sup>	.163 <sup>§</sup>	p <.000 <sup>§</sup>	p <.000 <sup>§</sup>
LEP op leg	p <.000 <sup>§</sup>	p <.000 <sup>§</sup>	N/A		.009 <sup>§</sup>	.002 <sup>§</sup>	p <.000 <sup>§</sup>	p <.000 <sup>§</sup>
Timed Walk	.012 <sup>§</sup>	.004 <sup>§</sup>	N/A		.009 <sup>§</sup>	<i>p &lt; .000<sup>§</sup></i>	p <.000 <sup>§</sup>	p <.000 <sup>§</sup>
Timed Sit to Stand	.001 <sup>§</sup>	.001 <sup>§</sup>	N/A		.672 <sup>§</sup>	.154 <sup>§</sup>	.005 <sup>§</sup>	p <.000 <sup>§</sup>

The change scores in means or medians, as appropriate, are provided in the tables below.

**Table 17. Table presenting the primary outcome Oxford knee score median change scores within the two groups for all time points.**

Outcome	Baseline and 3/12 Mean/ median change (Difference between gps)		3/12 and 6/12 Mean/ median change (Difference between gps)		6/12 and 12/12 Mean/ median change (Difference between gps)		Baseline and 6/12 Mean/ median change (Difference between gps)		Baseline and 12/12 mean/ median change (Difference between gps)	
	Home n=45	Usual n=46	Home n=40	Usual n=42	Home n=42	Usual n=44	Home n=41	Usual n=43	Home n=46	Usual n=47
Median Oxford Knee Score	11 (3)	14	3 (1.5)	1.5	2 (1)	1	15 (1)	16	16 (1)	17
95% Confidence Interval * lower, upper	8, 14.5	10, 16.5	0, 4	-5, 4	.5, 4.5	0, 4	12, 18	12, 18.5	14, 20	13.5, 20

\* for the median

**Table 18. Table presenting the secondary outcome mean / median change scores within the two groups for all time points.**

Outcome	Baseline and 3/12 Mean/ median change (Difference between gps)		3/12 and 6/12 Mean/ median change (Difference between gps)		6/12 and 12/12 Mean/ median change (Difference between gps)		Baseline and 6/12 Mean/ median change (Difference between gps)		Baseline and 12/12 mean/ median change (Difference between gps)	
	Home	Usual	Home	Usual	Home	Usual	Home	Usual	Home	Usual
Mean KOOS symptoms subscale	n=46 26.4 (0.14)	n=47 26.26	n=40	n=42	n=40	n=44	n=42	n=44	n=44	n=48
Median	32 (4)	28	7 (6)	1.5	3.86 (3)	6.86	36 (2.5)	33.5	39.36 (3.57)	35.79
Median KOOS pain subscale	n=46 28 (0)	n=47 28	n=40 3.5 (0.5)	n=41 3	n=40 2.94 (2.39)	n=43 5.33	n=42 34.5 (1.5)	n=43 36	n=44 40.56 (6.38)	n=48 46.94
Mean KOOS ADL subscale	n=40 24.5 (4.74)	n=41 29.24	n=35	n=37	n=36	n=40	n=39	n=39	n=39	n=43
Median	25.5 (4.5)	30	5 (5)	0	2.71 (3.79)	6.5	31 (3)	28	33.53 (1.47)	35
Median KOOS sport & recreation subscale	n=17 25 (5)	n=18 20	n=14 11.5 (11.5)	n=19 0	n=23 5 (0)	n=21 5	n=24 43.5 (28.5)	n=23 15	n=26 45 (17.5)	n=18 27.5
Median KOOS quality of life subscale	n=46 31 (0)	n=46 31	n=40 0 (0)	n=42 0	n=40 6.13 (.13)	n=44 6	n=42 32 (6)	n=42 38	N=44 40.63 (3.63)	n=47 37
Median leg extensor press of maximum W/kg for operated leg	n=39 .28 (.005)	n=38 .275	N/A		N/A		N/A		n=36 .41 (.015)	n=40 .395
Median leg extensor press of maximum W/kg for non operated leg	n=41 .130 (0.06)	n=40 .1900	N/A		N/A		N/A		n=36 .205 (.01)	n=42 .195
Median timed walk test in seconds	n=42 -1.34 (.47)	n=43 -.87	N/A		N/A		N/A		n=40 -1.93 (0.01)	n=43 -1.94
Median number of completed sit to stands in 30 seconds	n=42 2 (1)	n=43 1	N/A		N/A		N/A		n=41 3 (1)	n=43 2



For the Oxford Knee Score it can be seen that there is only a one point difference between the two groups baseline to 12 months scores. There is a 3 point difference, in favour of the usual physiotherapy care group, at the baseline to 3 months period.

For the KOOS, there are no large differences between the groups except for the sports and recreation subscale. Here there are differences of up 28.5 in favour of the home visit physiotherapy group. However, as the table shows, completion of this subscale was low. The majority of participants did not complete this subscale at baseline and three months and many still found it inappropriate at 6 and 12 months. The two groups appeared similar for all other outcomes at all other time points.

#### **4.7.5 Between group comparisons of secondary outcome follow up data.**

Outcome scores between each group were compared between each timepoint. Independent samples t-tests were used for the outcomes comparing normally distributed data and Mann-Whitney U tests for outcomes comparing non normally distributed data. The results can be seen in the table below.

**Table 19. Table presenting the comparisons of scores between groups at baseline, 3, 6, and 12/12 timepoints.**

Outcome	3/12 FOLLOW UP P values Independent t-test <sup>#</sup> OR Mann-Whitney U <sup>§</sup>	6/12 FOLLOW UP P values Independent t-test <sup>#</sup> OR Mann-Whitney U <sup>§</sup>	12/12 FOLLOW UP P values Independent t-test <sup>#</sup> OR Mann-Whitney U <sup>§</sup>
Oxford Knee Score	.288 <sup>§</sup>	.681 <sup>§</sup>	.937 <sup>§</sup>
KOOS symptoms	.549 <sup>#</sup>	.539 <sup>§</sup>	.991 <sup>§</sup>
KOOS pain	.330 <sup>§</sup>	.334 <sup>#</sup>	.271 <sup>§</sup>
KOOS ADL	.330 <sup>#</sup>	.894 <sup>#</sup>	.577 <sup>§</sup>
KOOS sport & recreation	.739 <sup>#</sup>	.087 <sup>#</sup>	.894 <sup>#</sup>
KOOS QoL	.679 <sup>§</sup>	.834 <sup>#</sup>	.600 <sup>§</sup>
LEP op leg	.919 <sup>#</sup>	N/A	.765 <sup>§</sup>
Timed Walk	.589 <sup>§</sup>	N/A	.852 <sup>§</sup>
Timed Sit to Stand	.811 <sup>§</sup>	N/A	.302 <sup>§</sup>

As can be observed, there are no significant differences between the groups for any outcome at any timepoint.

#### **4.7.6 One-way Repeated Measures Analysis of Variance.**

The findings from the repeated measures tests are summarised below in the table and visually represented via profile plots in the graphs below. The graphs show the changes over time for both groups and show the lines for the two groups generally to be close together with the results indicating the variables to be non significant. The groups on the sports and recreation subscale and timed walk test graphs are the farthest apart; the differences in seconds for the latter at 3 and 12 months follow up are still small.

**Table 20. Table to show results from repeated measures tests for all outcomes.**

Outcome	Mauchly's test of sphericity	Levene's test of equality of error variances p values	Box's test of Equality of Covariance Matrices	Interaction Effect Wilk's Lambda p values	Main effect Wilk's Lambda p values	Between-subject tests P values
Oxford Knee Score	.001	Baseline .076 3/12 .494 6/12 .714 12/12 .874	.337	.986	p<.000	0.434
KOOS symptoms	p<.000	Baseline .056 3/12 .802 6/12 .137 12/12 .133	.128	.622	p<.000	.753
KOOS pain	p<.000	Baseline .734 3/12 .361 6/12 .195 12/12 .535	.736	.822	p<.000	.336
KOOS ADL	p<.000	Baseline .054 3/12 .228 6/12 .255 12/12 .283	.845	.490	p<.000	.713
KOOS sport & recreation	.206	Baseline .022 3/12 .004 6/12 .327 12/12 .103	.483	.221	p<.000	.614
KOOS QoL	.013	Baseline .793 3/12 .280 6/12 .465 12/12 .503	.444	.855	p<.000	.720
LEP op leg	.002	Baseline .248 3/12 .861 12/12 .936	.710	.8	p<.000	.796
Timed Walk	p<.000	Baseline .052 3/12 .098 12/12 .032	p<.000	.545	.002	.256
Timed Sit to Stand	.025	Baseline .433 3/12 .139 12/12 .712	.231	.438	p<.000	.510

Estimated Marginal Means of MEASURE\_1

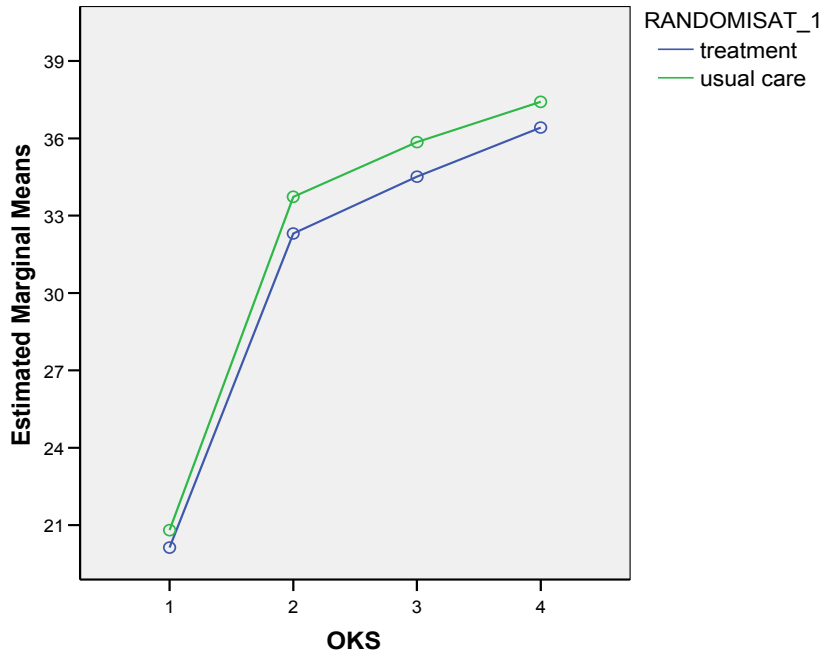


Figure 16. SPSS graph to show the one-way repeated measures analysis of variance for the Oxford Knee Score across all timepoints for both groups.

Estimated Marginal Means of MEASURE\_1

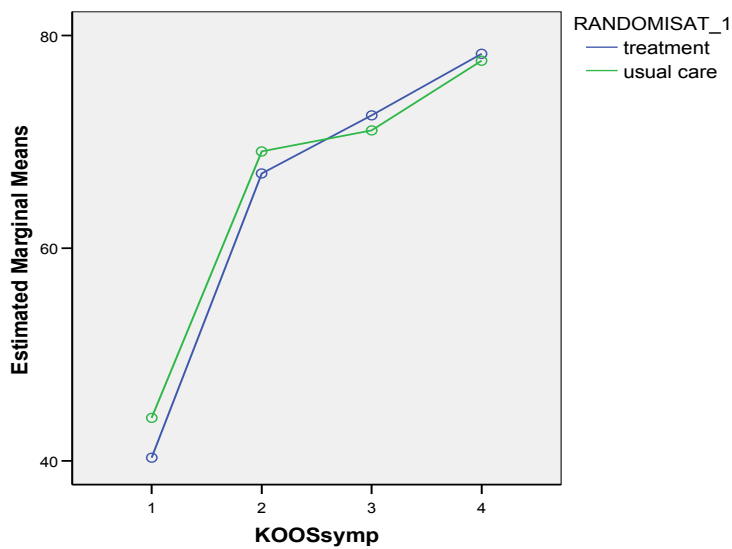
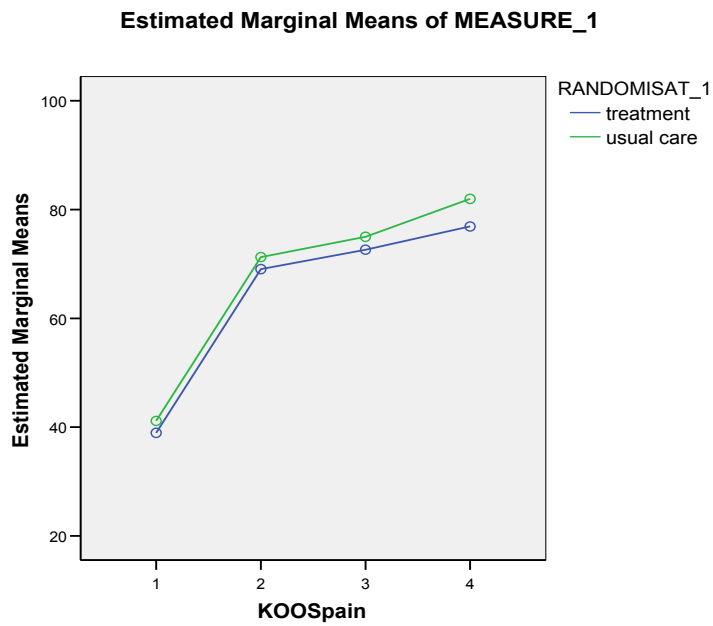
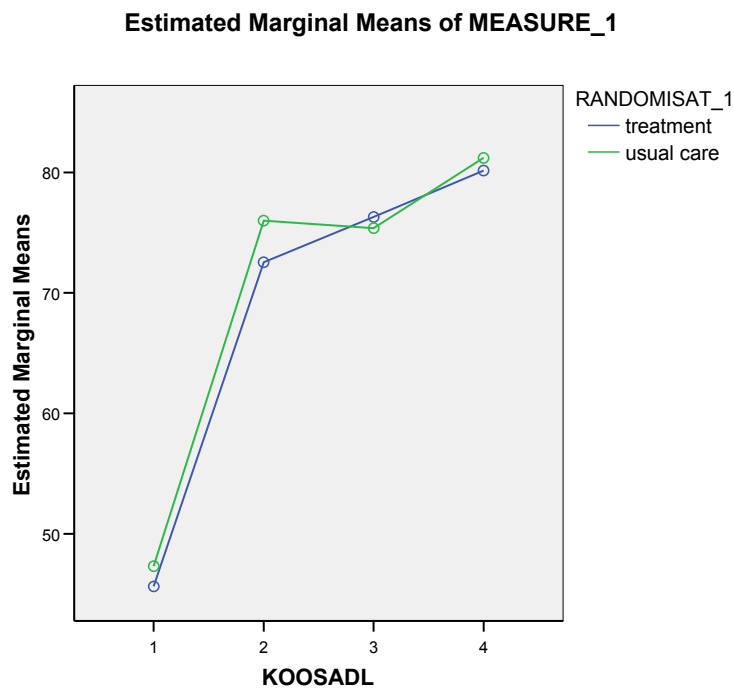


Figure 17. SPSS graph to show the one-way repeated measures analysis of variance for the KOOS Symptoms subscale across all timepoints for both groups.

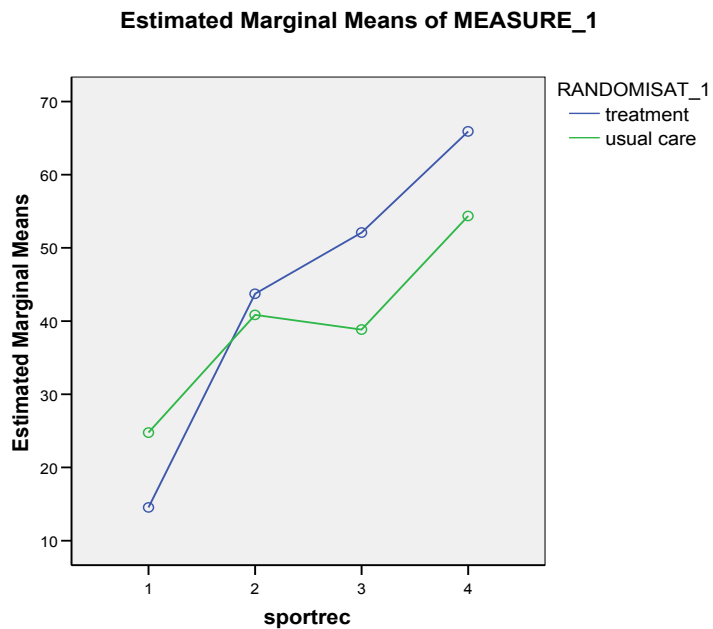


4a

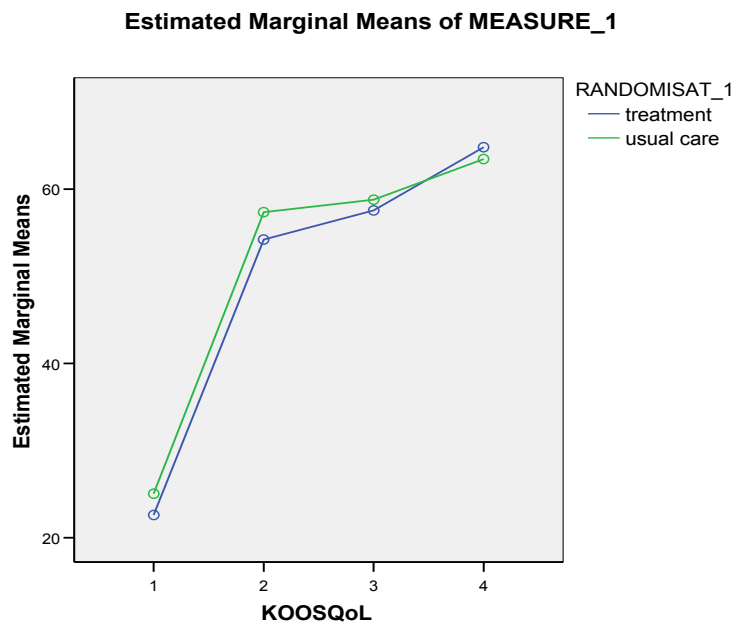
**Figure 18.** SPSS graph to show the one-way repeated measures analysis of variance for the KOOS Pain subscale across all timepoints for both groups.



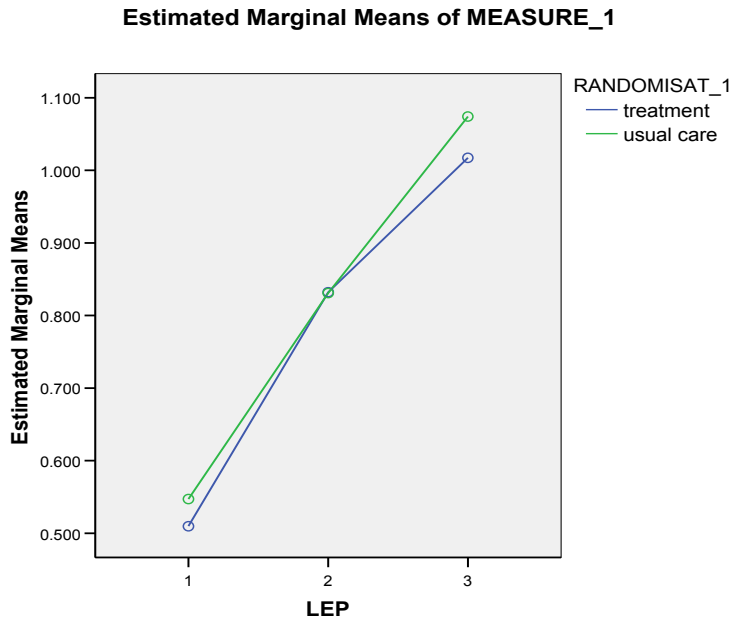
**Figure 19.** SPSS graph to show the one-way repeated measures analysis of variance for the KOOS ADL subscale across all timepoints for both groups.



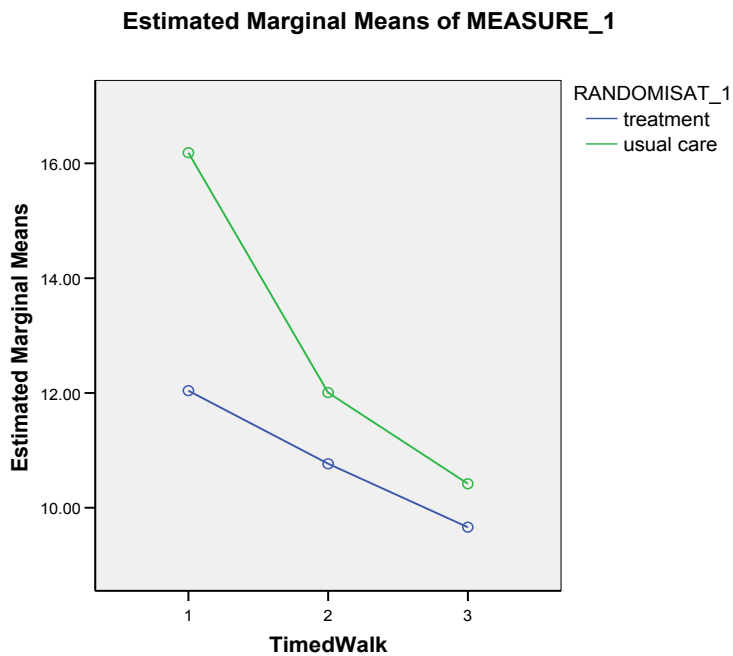
**Figure 20.** SPSS graph to show the one-way repeated measures analysis of variance for the KOOS Sports and Recreation subscale across all timepoints for both groups.



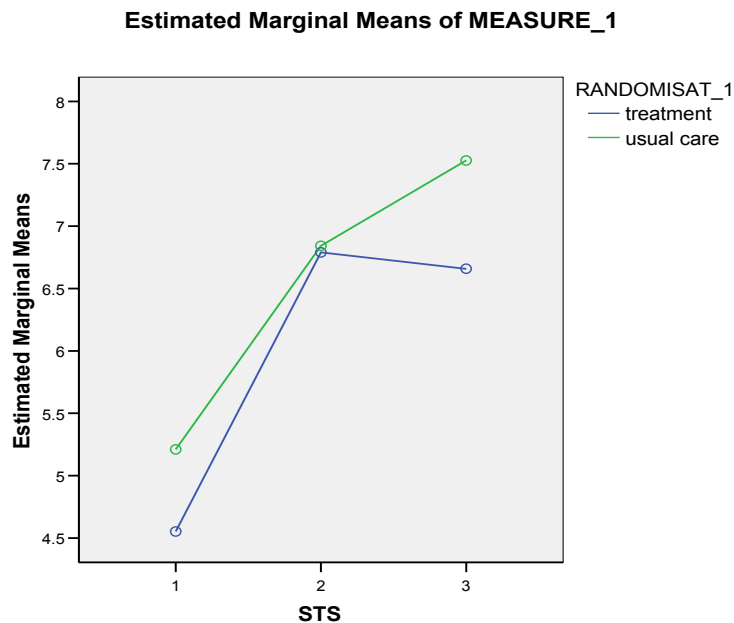
**Figure 21.** SPSS graph to show the one-way repeated measures analysis of variance for the KOOS Quality of Life subscale across all timepoints for both groups.



**Figure 22.** Graph to show the one-way repeated measures analysis of variance for the Leg Extensor Press (operated leg) subscale across baseline, 3 and 12 month timepoints for both groups.



**Figure 23.** Graph to show the one-way repeated measures analysis of variance for the Timed Walk test across across baseline, 3 and 12 month timepoints for both groups.



**Figure 24. Graph to show the one-way repeated measures analysis of variance for the Timed Walk test across across baseline, 3 and 12 month timepoints for both groups.**

Mauchly's tests for sphericity were significant for all outcomes bar the KOOS sports and recreation subscale. For significant results it is assumed that there are significant differences between the variances of differences and thus the subsequent *F*-ratios produced become dubious (Field, 2000 p.323-6). Although violations can be compensated it is considered more sound to inspect the multivariate statistics instead, as follows (Pallant, 2007 p.272).

Levene's test of equality of error variances were checked to see if the assumption of homogeneity of variances were violated (Pallant, 2007 p.272). No outcomes had significant values for all timepoints and most were non significant across all timepoints. It is concluded that this assumption has not been violated.



Box's test of equality of covariance matrices then tested the null hypothesis that the observed covariance matrices of the dependent variables are equal across groups (Pallant, 2007 p.272). All results were not significant bar timed walk where this assumption is violated.

The one-way repeated measures analyses of variance showed that there were no significant interactions between group type and timepoint. Interaction effects were non statistically significant throughout. There were however significant effects for time across both groups and for all outcomes. All tests of between-subject effects were non significant so it is concluded that the main effect for group is not significant. The lack of sphericity however lessens the dependability of these results.

#### **4.7.7 Range of Knee Joint Motion.**

The data regarding knee joint range of motion is summarised below. As can be seen the groups appear similar at each timepoint except that the home visit group have slightly more flexion than the usual care group.

**Table 21. Table to show average knee joint range of motion in degrees for operated leg at baseline, 3 and 12 months timepoints**

Group	Baseline Extension Flexion		3 Months Extension Flexion		12 Months Extension Flexion	
Home Visit Physiotherapy	FFD 9.3	105.3	FFD 6.39	105	lacks 3.83	108.4
Usual Physiotherapy	FFD 9.9	105.4	FFD 6.1	102.9	lacks 4.64	107.6

FFD = fixed flexion deformity (i.e. >5 degrees lack).

#### **4.7.8 THE PHYSIOTHERAPY INTERVENTION.**

Every visit to trial participants receiving the trial intervention was recorded on a case report form (Appendix XIII). 49 participants received visit one and 48 received visit two (the remaining participant withdrew from the trial before visit two stating she was doing very well and did not need any further treatment). The content and estimated times (in minutes) spent on each treatment component were detailed. The tables below present the summarised data for each of the two trial visits. Apart from the duration of visit one, the rest of the data were not normally distributed (see Appendix XV) so the medians are presented rather than the means. The most time intensive component was the general subjective and objective assessment of patients and progression of walking aids. The median times spent on the specific components of the intervention were 10-11 minutes in visit one and 8-13 minutes in visit two. The individualised components of treatment, linked to activity restrictions and participation restrictions were generally included in the task specific training summary time since they included practicing/altering activities. Examples of such activities include using public transport, cleaning rayburns and getting in and out of armoured vehicles.

**Table 22. Summary of the content and time (in minutes) of the home visit treatments provided in the trial intervention.**

Content	Visit One: within 2/52 of discharge		Visit Two: 6-8 weeks after discharge	
		no of participants with data		no of participants with data
<b>Duration of Visit</b>				
Median (Interquartile range)	45 (15)	n=49	40 (31)	n=46
Range	25-70		20-79	
<b>General</b> (subjective and objective reassessment)				
Median (Interquartile range)	19 (12)	n=49	14 (14)	n=47
Range	9-42		6-40	
<b>Exercises</b> (Teaching, checking, progressing)				
Median (Interquartile range)	10 (7)	n=49	8 (4)	n=47
Range	5-22		(4-18)	
<b>Task Specific Training</b> (car, rising from chair, walking outside, stairs)				
Median (Interquartile range)	11 (5)	n=49	13 (7)	n=47
Range	2-23		5-30	
<b>Manual Therapy</b>				
Median (Interquartile range)	0 (5)	n=49	0 (0)	n=47
Range	(0-9)		0-7	
<b>Other</b> (eg. rest, listening)				
Median (Interquartile range)	2 (5)	(n=49)	1 (5)	n=48
Range	0-18		0-23	

The two physiotherapists who carried out the home visits also listed key concerns raised by patients during home visits and these are summarised in the table below. Thirty five participants raised concerns during their first home visit and 28 during visit two. The main concerns in visit one were complications or issues relating to the knee operation (19/66 or 29%), rate of recovery (14/66 or 21%) and range of knee joint motion (10/66 or 15%). By visit two these concerns were expressed less frequently and the most frequent concern had shifted to returning to activities of daily living (12/39 or 31%).

**Table 23. Summary of the concerns raised by participants and recorded by physiotherapists during the home visit treatments provided in the trial intervention.**

<b>Concern</b>	<b>Visit One: within 2/52 of discharge (no of participants raising concern)</b>	<b>Visit Two: 6-8 weeks after discharge (no of participants raising concern)</b>
Complications:	swelling n=6 numbness n=1 DVT n=1 calf muscle soreness n=1 open wound n=4 itchiness n=1 cellulitis n=1 bruising n=2 stiffness n=1 redness n=1	swelling n=1 stiffness n=1 scar tissue n=1 heat in knee n=1 open wound n=1 calf pain n=1 depression n=1
Fatigue	2	
Pain and aching	3	4
Range of knee joint motion	10	5
Knee instability		1
Rate of recovery	overall rate of recovery n=8 how much weight bearing n=2 walking aid progression n=3 overdoing it n=1	1 1
Lack of Confidence in knee	3	1
Activities	return to driving n=1 stairs n=1 shower n = 1 flying n=1 using bus n=1 kneeling n=1 return to work n=1	return to driving n=1 cycling n=1 swimming n=1 bath transfers n=1 kneeling n=7 return to work n=1
Poor balance and fear of falling	1	1
Question regarding operative procedure	2	
Co-morbidities	other knee n=1 shoulder pain n=1 lymphoedema n=1 other mobility problems n=1	other knee pain n=1 other joint pain n=1 varicose veins n=1 law back pain since op n=1
Medication query	1	
Awaiting hospital follow up		2
<b>Total number of concerns</b>	<b>66</b>	<b>39</b>

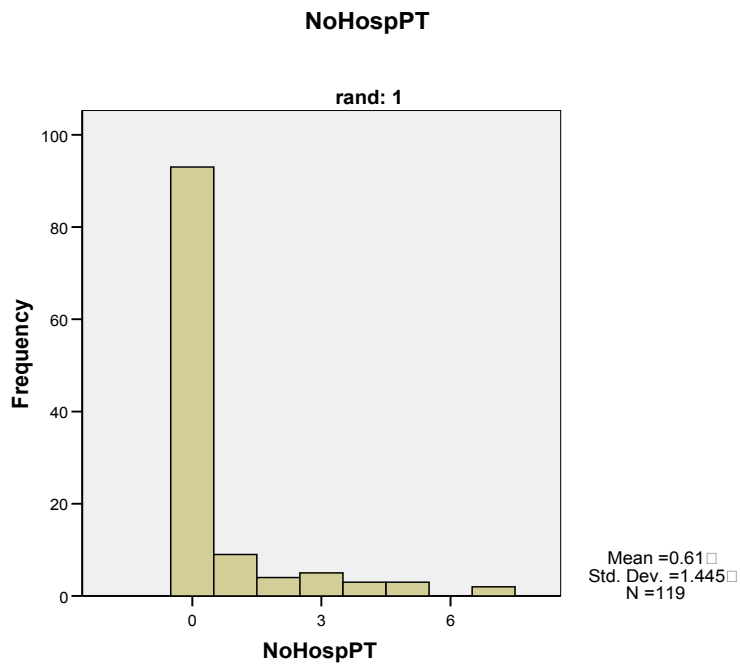
#### **4.7.9 ADDITIONAL PHYSIOTHERAPY PROVIDED TO PARTICIPANTS DURING THE TRIAL.**

The patient resource diaries requested information regarding all non trial physiotherapy contacts from participants because of the impact this could have upon trial results. This information is summarised in the table and histograms below. More physiotherapy contacts (n=119) were recorded by the usual physiotherapy care group than in the home visit treatment group (n=93). The majority of visits (157 or 74.06%) occurred in the first three months, with 29 (13.68%) in 3-6 months and 26 (12.26%) for 6-12 months post operatively.

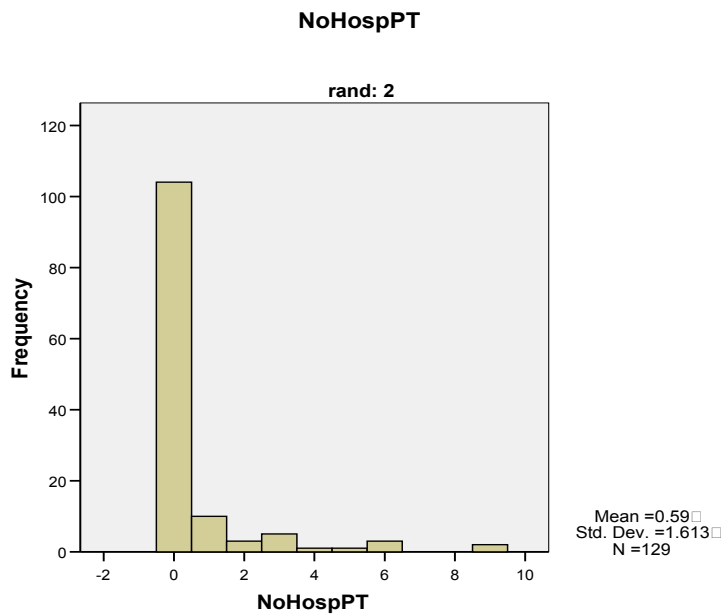
**Table 24. Table to show the total number of non trial physiotherapy visits occurring between discharge from hospital after their knee replacement and one year follow up.**

	Home visit group (n=144 diaries) No of visits in group (no and % of diaries)	Usual physiotherapy group (n=147 diaries) No of visits (no and % of diaries)
Number of hospital outpatient visits	73 visits (n=26; 18.06%)	76 (n=25; 17.01%)
Number of general practice visits	13 visits (n=2; 1.39%)	15 (n=6; 4.08%)
Number of home visits	1 (n=1; 0.69%)	0
Number of private health care visits	6 (n=2; 1.39%)	5 (n=2; 1.36%)
Number of visits during nursing home care		9 (n=1; 0.68%)
Police intensive rehabilitation programme		7 full days* (n=1; 0.68)
Total number of physiotherapy contacts	93	119

A conservative estimate of two sessions a day has been used in the calculation of the total number of physiotherapy contacts.



**Figure 25. Histogram showing the number of outpatient physiotherapy visits for the home visit physiotherapy group.**



**Figure 26. Histogram showing the number of outpatient physiotherapy visits for the usual physiotherapy group.**

It became apparent that participants were recording their research trial and follow up visits. All trial visit dates were checked against the diary data allowing these visits to be removed prior to analysis. Five participants wrote that some of their outpatient hospital appointments were for hydrotherapy. Hydrotherapy visits ranged from 1-6 and totalled 16 in total; all five participants were in the usual physiotherapy group. It also became apparent that at least three of the hospital outpatient physiotherapy appointments recorded by participants (intervention n=1 and usual physiotherapy n= 2) were for the orthopaedic outpatient discharge clinic rather than for physiotherapy treatment.

#### **4.7.10 SAMPLE SIZE.**

As so little information regarding sample size previously existed Altman's nomogram to calculate sample size or study power was used to estimate the power of this study and to produce a power calculation for future work from the data this trial produced (Altman, 1991 p. 456-457).

##### **4.7.10.1 Power.**

For a clinical difference of 2.5 points between groups on the Oxford knee score the standardised difference is as follows:

$$\begin{aligned} 2.5 \text{ divided by } 1 \text{ standard deviation} &= 2.5 \text{ divided by } 7.857 \\ &= 0.318 \end{aligned}$$

For a trial size of 100, the nomogram estimates the power of the present study to be 35%.

##### **4.7.10.2 Sample Size.**

For a standardised difference of 0.318, 90% power and a significance level of 0.05 a sample size of 400, or 200 in each treatment arm, would be required.

## **4.8 DISCUSSION.**

### **4.8.1 SECTION OVERVIEW.**

This chapter contains the discussion of the methodology and findings of the trial comparing post discharge physiotherapy versus usual care following primary total knee arthroplasty for osteoarthritis. The chapter begins with a summary of the main findings of the trial. It then provides a discussion of the strengths and limitations of the trial's procedures followed by a discussion of how this trial compares to previous trials. Following this the clinical implications of the study and directions for future research are presented.

### **4.8.2 SUMMARY OF MAIN FINDINGS.**

The alternate hypotheses for the trial, that participants receiving two home physiotherapy visits after discharge from hospital would show greater improvement in outcomes than those participants were receiving usual physiotherapy, were rejected. The null hypotheses, that there would be no statistical differences in outcomes between the two groups, were accepted.

#### **4.8.2.1 Between group differences.**

There were no significant differences observed between the home visit physiotherapy group and the usual physiotherapy group for any outcome at any time point in this phase II trial. However, as will be discussed, the trial was under powered which, when considering the small treatment effect size expected from an intervention like physiotherapy, may have meant the trial was too small to have picked up any treatment effect which might have existed. In addition, important factors such as the development of major co-morbidities and further lower limb surgery during the trial, were seen to affect the home visit physiotherapy group more than the usual physiotherapy group which could have obscured group differences in this trial.



#### **4.8.2.2 Within group differences.**

The baseline to three months scores within group scores are significantly different for all outcomes in both groups. The values for both groups appear similar.

At the 3-6 months time points the KOOS symptom subscale showed a significant difference within the home visit group but not within the usual care group. The KOOS ADL and sport and recreation subscales values for each group also appeared dissimilar.

At the 6-12 months time points there was a significant difference in Oxford knee scores within the home visit group but not within the usual physiotherapy group. Both groups showed a significant difference in scores for the KOOS symptoms. The usual physiotherapy group shows significant differences in KOOS pain, functional activities of daily living and sports and recreational scores whereas the home visit group did not.

At the 3-12 months time points, for outcomes not collected at 6 months, both groups showed significant differences for the Leg extensor press scores and the timed walk scores and both showed non significant differences for the time sit to stand scores.

#### **4.8.2.3 Repeated measures.**

The one-way repeated measures analyses of variance showed that there were no significant interactions between group type and timepoint. Interaction effects were non statistically significant throughout. There were however significant effects for time across both groups and for all outcomes. All tests of between-subject effects were non significant so it is concluded that the main effect for group is not significant. The lack of sphericity however lessens the dependability of these results.

### **4.8.3 STRENGTHS AND LIMITATIONS OF THIS RESEARCH.**

#### **4.8.3.1 Trial procedures.**

The rigorous and comprehensive trial procedures used in this trial are believed to have added to its quality, allowing a high standard of reporting. The provision of additional procedures and information regarding issues influencing the internal validity of the trial, which are often left unreported in the space restrictions of trial reports, enable the trial to be placed into context more meaningfully than any previously published similar physiotherapy after knee joint arthroplasty trial. These procedures will now be discussed.

#### **4.8.3.2 Randomisation and allocation concealment.**

As can be seen from the results, the two trial arms appeared relatively similar in terms of their baseline characteristics. The randomisation processes used in the trial therefore appeared successful. However, as will be seen in the complications section below, this did not mean that the groups remained similar after randomisation. Important differences, believed to be unrelated to the intervention, did subsequently arise between the two groups.

The mean ages and standard deviations of the two groups reflect the most common age groups undergoing total knee arthroplasty according to the National Joint Registry 4<sup>th</sup> and 5<sup>th</sup> Annual Reports. Whilst age affects likelihood of revision, with younger patients being more likely to undergo revision, the existing evidence does not suggest that age affects functional outcome (Santaguida *et al.*, 2008; Dinah and Mears, 2008).

The functional co-morbidity mean at baseline appears low especially when, by the nature of their having osteoarthritis, all participants had to score at least one (Groll *et al.*, 2005). A greater number of comorbid conditions is associated with worse short-term pain

relief and poorer functional outcome post operatively (Jones *et al.*, 2007). Solomon *et al.*, (2006) report that 36.7% of patients in their study have at least one co-morbid condition, less than the current trial, although comorbidity is defined differently using an index for use with ICD codes. Roos and Toksvig-Larsen (2003) report an average similar to this trial of 1.3 co-morbid conditions amongst their total knee arthroplasty sample. The mean BMI scores at baseline were in the obese category for the home visit group and close to the obese category for the usual physiotherapy group. Since obesity is known to be a significant risk factor for osteoarthritis these means are considered usual for this patient group, and are even slightly lower than the mean of 32.6 in patients undergoing primary total knee joint arthroplasty found in a recent retrospective review by Fehring *et al.* (Hunter and Felson, 2006; Fehring *et al.*, 2007).

Few participants in this trial reported they smoked (n=6) therefore smoking, known to influence outcome and complication rates following lower limb arthroplasty (Cowie *et al.*, 2009; Møller *et al.*, 2002), was unlikely to have impacted upon the findings of this trial. Less than 10% of participants had diabetes mellitus, a condition increasing the risk of complications such as stroke and pneumonia (Bolognesi *et al.*, 2008). The number of patients reporting a cardiac history was higher at 28.6%. No attempt was made to confirm more precise diagnoses, so this figure may be an overestimate. Cardiac problems are a risk factor for cardiovascular complications following joint arthroplasty (Basilico *et al.*, 2008) but, as will be seen in the complications section below, there were no significant differences between the groups with regard to cardiac complications.

The knee joint ranges of motion for the operated leg were similar for both groups. Both groups reduced their average fixed flexion deformities from 9-10 degrees at baseline to 3.8-4.6 degrees at twelve months post operatively. This range still exceeds the normal range

of knee extension, extension limitations of two degrees are considered normal for adults, and is close to five degrees which is the range at which limitations in extension become a fixed flexion contracture (Norkin and White, 2003 p.224). Knee flexion increases slightly from 105 at baseline to 107-108 at twelve months. This is slightly above the 105 degrees of optimal range of knee flexion postoperatively (Jones *et al.*, 2007).

There were no instances of participants knowing their allocation, or thus being able to reveal this to the physiotherapist undertaking the baseline assessment, prior to the baseline assessment being completed. All allocation envelopes were used in their correct sequence. Allocation concealment from assessing physiotherapists and participants was therefore maintained until after the baseline assessment as intended. Although the participants then needed to be aware of their group allocation a high level of outcome assessment blinding was recorded for the trial; this is the subject of the next chapter and will be discussed in depth there.

#### **4.8.3.3 Sample size.**

The results show the trial to be underpowered. Altman's nomogram to calculate sample size (Altman, p. 456-7) estimates that a sample size of 400, 200 in each treatment arm, would be required to achieve a study with 90% power. The current trial suggests that this number would really need to be substantially greater. To allow for attrition and drop out, to allow for those participants who receive a uni compartmental arthroplasty rather than a total and to allow for participants subsequently deciding to have a second arthroplasty within the follow up period the sample size would need to be increased by at least 20%, or an additional 40 participants in each arm.

#### **4.8.3.4 Complications rates.**

There is a general lack of data regarding postoperative complications following total knee joint arthroplasty in the UK. This trial provides self report data from trial participants regarding major and minor complications. Major complications occurring within the first three months were then checked against patients' medical notes. The trial is therefore able to report more fully regarding complications than any previous physiotherapy trial following knee joint replacement. Before commenting upon the complication rates from this trial, I would like to mention several reflections regarding complication rate measurement which this trial raised. It became apparent that patients report what they perceive to be important to them and their recovery. Many of their concerns were also repeated/reflected in the concerns raised by participants during treatments and listed in the case report treatment forms. Such perceptions may and did vary from medical opinions. Many of the medically considered "minor" complications were given great import by patients. The opposite was also seen when one participant did not mention a medically considered important deep vein thrombosis.

There was no one accurate source of post operative complications in this trial. The hospital notes often did not include serious complications developed by patients. For example two out of three patients reporting the development of pressure sores did not have this mentioned in their hospital notes; supporting evidence for the existence of these sores was obtained via the resource diaries listing multiple district nurse visits for pressure sore dressing. Also, complications arising in primary care after discharge did not always get recorded in medical notes and it may be that surgeons remain unaware of them.

So the question of what complications to measure, and how to measure complication rates thus becomes very interesting. Should we be measuring and reporting the wide range of complications patients believe to be relevant to their recovery or should we measure the post

surgical complications surgeons are interested in? Surely there is need for both? Where is the best source of information to be found? Should we examine primary care records as well as hospital ones?

Within this small trial the complication rates for those complications for which existing data is available were generally high. This is despite studies suggesting that patients managed at specialist orthopaedic hospitals by surgeons with high arthroplasty caseloads have lower risks of perioperative adverse events (Katz *et al.*, 2004; Katz *et al.*, 2008). The revision rates within one year of surgery for England and Wales (National Joint Registry 5<sup>th</sup> Annual Report) is 0.3%. In this trial the revision rate was 2.1% (n=2/93). This may be high due to the one participant obtaining an intra-operative fracture and requiring revision. Such fractures are rare; there were only 102 in 2007 in England and Wales out of 61,389 patients on the joint registry. The increase in complication rates may be by chance, or may be partially explained by the fact that even so called straightforward knee joint arthroplasties occurring at a specialist orthopaedic centre may transpire to be more problematic referrals than the national norm.

Not all complication rates were above average. The National Joint Registry 5<sup>th</sup> Annual Report reports a 0.5% mortality rate at three months post operation. In this trial no patients died within the first three months. The 4<sup>th</sup> Annual Report reports the mortality rates within one year of surgery for England and Wales to be 1.6%. For this trial the mortality rate was 1.03% (n=1/97) for this time period.

Although a recent review examining post operative complications in the over 70s age group concludes there is a lack of research available (Dinah and Mears, 2008) there is some additional information regarding complication rates from outside the UK, mainly from the US. Parvazi *et al.*, (2007) report a 0.37% rate of stroke following knee and hip replacement

during hospitalisation and the first 6 weeks post discharge; in the current small trial the rate was much higher (3.3% in the first three months and then 2.17% until one year) although both stroke and transischaemic attacks were included in the current trial.

Previously reported superficial wound infection rates range from 0.26% in the first three months (Peersman *et al.*, 2001) to 10.5% post knee replacement (Gaine *et al.*, 2000). The Scottish authors of the latter study say their higher figure is due to their encompassing grading system. The current trial rate is even higher at 16.3%; these figures are self report figures rather than healthcare records and may also be raised by antibiotics being used prophylactically, when a superficial infection is suspected, rather than following confirmed bacterial growth. For deep infections, Solomon *et al.*, (2006) report a 0.3% rate in the first 90 days and a 0.89% rate in the first year has just been reported for Finland (Jämsen *et al.*, 2009) ; the lack of agreement between patient self-report and the medical notes make it hard to calculate a rate for the current trial. Soohoo *et al.*, (2006) report an infection rate of 0.71% although the level of infection is not clear.

For pulmonary emboli Solomon *et al.*, (2006) report a 0.9% rate in the first 90 days post operation, similar to the current trial value of 1.1%. Soohoo *et al.*, (2006) report a lower 0.43% for the same time period.

Solomon *et al.*, (2006) provide an adverse event rate for the first 90 days post operation (including pulmonary emboli, myocardial infarction, pneumonia, deep wound infection and death) as 3.6%. A similar adverse event rate for the current trial is hard to calculate, for example the chest infections may not have been tested for pneumonia, but would be somewhere between 2.09-7.26%. However, the narrowness of complications included in the rate reported by Solomon *et al.*, for example excluding stroke, can not be forgotten.

#### **4.8.3.5 Retention/follow up.**

As reported in the CONSORT flow diagram there was little loss to follow up in the trial. In the usual care group there was no loss to follow up at one year. The primary outcome was collected for 48/49 participants who underwent surgery (one participant died). In the home visit the primary outcome was collected for 47/49 participants who underwent surgery (one withdrawal and an Oxford knee score could not be calculated for one person who did not answer sufficient items). It has been shown previously that loss to follow up does matter for joint arthroplasty patients at this hospital Trust; over time patients lost to follow up have a worse outcome than those who continue to be assessed (Murray *et al.*, 1997). Reluctance to attend for follow up cannot be underestimated; loss to follow up at 12 months in one previous physiotherapy feasibility trial was nearly 50% (Frost *et al.*, 2002).

However, one striking follow up finding was that more patients would have been lost to the trial if the primary outcome measure had not been the self report Oxford knee score. As can be seen from the lower numbers of participants for whom secondary outcomes were obtained at one year, participants were less willing to attend for face-to-face follow up at the hospital and some postal and telephone follow up assessments were required to prevent withdrawals from the trial. Interestingly, as stated in the results, many patients stated reluctance to travel as their main reason for non participation in the trial. Preventing travel for follow up assessments may well prove to be a successful strategy in the recruitment and retention of patients therefore: the need for such strategies is recognised (Frost *et al.*, 2002).

#### **4.8.3.6 The Intervention.**

As stated in Chapter four, trials can be criticised as being too pragmatic if the interventions are described as being totally at the discretion of the health professional and if no clear



definitions of the interventions are provided (Koes and Hoving, 1998). Without standardising or defining the interventions precisely then the generalisability of the results of a trial cannot be adequately assessed. Adherence to the trial protocol is a further issue. If both the content and consistency in delivery of the intervention is lacking then the internal validity of the trial may be undermined (Mullis *et al.*, 2006). This trial is one of the few physiotherapy trials to use case report forms. In these forms the two physiotherapists providing the intervention estimated the treatment time spent upon each component of the intervention for each patient in each home visit. I assisted in the development of such case report forms for use in a previous trial where they were validated against video analyses of treatment and found to have moderate to very good agreement; kappas=0.45-0.82 (Mullis *et al.*, 2006). A slight tendency whereby the time spent on content in the case report forms was underestimated compared to the video analyses was observed. In the current trial the case reports showed that all trial participants received the intervention. An estimated average time of 11 minutes was spent on task specific training in visit one and 13 minutes in visit two. An estimated average of 10 minutes was spent on functional exercises in Visit one and 8 minutes in visit two. If similar to the previous trial, these times may also be slight underestimates. One explanation for the lack of difference found between the two groups may be dosage. The time spent on the intervention may have been insufficient to affect outcome during the follow up time period.

A further issue thought likely to influence the internal validity of the trial was regarding any additional physiotherapy trial participants might received during the trial. Patients were therefore asked to list the number and location of any additional physiotherapy they received in a resource diary. Again, this is one of few trials to try and measure this aspect of the trial. Nearly a fifth of participants did receive additional physiotherapy with a total of 212 visits being recorded. However, it became apparent whilst data entering that patients

were, against instructions, also listing their trial follow up research visits. These were removed before analysis. Since it is possible, however, that some people attended out patient physiotherapy on the same day they attended their trial follow up assessments this may have lead to an underestimate of physiotherapy treatments. It also became apparent that at least three of the hospital outpatient physiotherapy appointments recorded by participants (intervention n=1 and usual physiotherapy n= 2) were for the orthopaedic outpatient discharge clinic rather than for physiotherapy treatment and it is likely that this number is an underestimate. It is therefore not known precisely how many of the outpatient appointments were routine follow up checks and how many were specific physiotherapy treatments.

Also, the diaries only collected the number of physiotherapy visits and it would be interesting to know the content of such visits. For example none of the home visit group mentioned receiving hydrotherapy, unlike the usual physiotherapy group. The format of the resource diary therefore requires amendment before future use to improve the accuracy and usefulness of the data captured.

#### **4.8.4 Use of Outcome Measures.**

##### **4.8.4.1 Oxford Knee Score.**

The Oxford knee score was found, as Dunbar *et al.*, (2001) have found previously, to be easy for patients to complete. This ease of completion was reflected by the completeness of data for this outcome: as can be seen in the CONSORT flow diagram very little missing data occurred in this trial for this outcome with response rates higher than the 85.9 -90% previously seen (Dunbar *et al.*, 2001; Robertsson and Dunbar, 2001; Whitehouse *et al.*, 2005). The preoperative mean for the current trial, of 20.11 (SD 7.857), and mean at one

year follow up (36.82, SD 8.908) are slightly better than those cited by Murray *et al.*, (2007) of 18 pre operatively (SD 7.5) and 34.2 (SD 10) at one year.

One theoretical concern regarding the Oxford knee score was whether a short 12 item score designed to measure outcome post arthroplasty would be sufficiently sensitive to measure the effects of physiotherapy. In previous knee joint arthroplasty trials (Frost *et al.*, 2002; Mockford and Beverland, 2008) no between group differences have been found using the Oxford knee score. However, its sister score, the Oxford hip score, has been used successfully to demonstrate between group differences in a previous physiotherapy trial (Trudelle-Jackson and Smith, 2004) and lack of sensitivity was not anticipated to be a serious problem for the trial. The far more detailed KOOS was included in the trial to provide more detailed examination of the domains of pain, symptoms and function and to supplement the Oxford knee score if required. Since both the detailed KOOS and the Oxford knee score showed no significant differences between groups at all time points it is believed that lack of sensitivity within the Oxford knee score did not adversely effect the findings of this trial.

During the time of this study several of the known issues with the Oxford knee score lead to a slight revision of the score. The wording of item 4 (walking) and item 7 (kneeling) had been criticised as being ambiguous (item 4) and inappropriate (item 7) (Whitehouse *et al.*, 2005). Susequently the score was slightly revised in response (Murray *et al.*, 2007). Item 4 now clarifies the “not at all” response by adding “pain severe when walking” to prevent patients from inferring the opposite meaning intended (of no pain). For items, such as item 7 (could you kneel), the word “could” is now in bold print to alert patients who have been told not to kneel that they need to fill this in hypothetically (Murray *et al.*, 2007). Although not believed to have significantly affected the study findings, it is accepted that the

use of the original score may have created some difficulties for patients when completing the score.

During the time of this study the score developers also published a recommendation that the minimum number of patients required in trials using the Score as the primary outcome should be at least 100 and usually many more (Murray *et al.*, 2007). The score developers aim to produce MCID estimates (minimal clinically important difference estimates, i.e. the smallest change in scores which patients perceive, and clinicians use, as a meaningful change) but suspect that such differences may be lower than 3 points on the Oxford knee score. Another indication, as subsequently shown, that the current trial is underpowered.

A further development since this study took place has been the publication of new research measures designed, unlike the Oxford knee score, to reflect all three of the main ICF components of impairment, activity limitation and participation restriction (Pollard *et al.*, 2009). Had measures like these been available at the time the trial took place it is believed these would have been, with the new osteoarthritis pain measure already mentioned (Hawker *et al.*, 2008), more appropriate outcomes for this trial. Research since the trial started has shown that the Oxford knee score mainly measures impairment and does not measure participation restriction (Pollard *et al.*, 2006). As participation restriction was included in the trial it is acknowledged that the Oxford knee score's inability to measure this means the score is less likely to have shown an intervention treatment effect in this trial.

One final point about the Oxford knee score data is that a normal distribution of data was seen at baseline. A similar finding, using the WOMAC score rather than the Oxford hip score, has recently been observed for patients awaiting total hip replacement (Dieppe *et al.*, 2009). This normal distribution suggests that many patients reported mild functional disability before having their knee replaced, even though, as in the Dieppe *et al.*, study, their X-ray

findings showed severe radiological changes. Dieppe *et al.*, thus comment upon the complexity involved in decision making regarding arthroplasty and when to perform surgery; likely to also include factors such as social circumstances, psychological status and motivation.

#### **4.8.4.2 KOOS.**

Subjectively it was noted that many participants balked and verbally commented regarding the length of the KOOS. Despite this the score was generally well completed except for the sports and recreation subscale which less than half the participants completed at most time timepoints. Participants said that the questions were inappropriate. These missing data rates initially appear hugely in excess of the 3.2% of total knee arthroplasty participants claimed by the scores developers (Roos and Lohmander, 2003). On closer examination however the reported 3.2 % missing data rate excludes the sports and recreation subscale and relates only to the remaining four subscales. As in this study, sports and recreation subscores could not be calculated for over half the participants (n= 58/105) used (Roos and Toksvig-Larsen, 2003). The authors argued that whilst irrelevant to many, the sports and recreation subscale remains useful by evaluating functions considered important to every second patient undergoing arthroplasty. In the current trial this subscale did show some interesting within group differences, for example there was a significant difference for the usual physiotherapy group at 6-12 months but not for the home physiotherapy group, but with so much missing data the group sizes become small.

The baseline KOOS and change subscale scores for the current trial and the only earlier knee joint arthroplasty trial reporting the KOOS (Roos and Toksvig-Larsen, 2003) were examined. Differences in means between the two trials varied by around 1-7 except for the sports and recreation subscale means at twelve months where the current trial's mean was

56.98 (Appendix XV) and was 46 in the trial by Roos and Toksvig-Larsen (2003). It has been suggested that 8-10 points represent minimal perceptible clinical improvement for KOOS subscales (Roos and Lohmander, 2003). The 10 point difference in means from the two different trials is therefore interesting; once again however, whether patients in the current trial did improve more in this aspect is difficult to ascertain given the amount of missing data and underpowering of both trials. For similar reasons it is difficult to judge whether the value of using the more detailed KOOS in physiotherapy trials, in addition to using the quick Oxford knee score, outweighs the time taken for its completion.

#### **4.8.4.3 Leg Extensor Press.**

The results show that not all participants were able or willing to undergo leg extensor press measurement. Pain was the predominant reason for being unable to do the test. The outcome assessment physiotherapists also noted that some participants appeared/stated they were afraid of the test despite verbal reassurance and the warm up test practices.

When compared against the results of the Allied Dunbar National Fitness Survey (1990) the trial results for the operated leg (ranging from median values .52 to .88 W/kg) show that the participants are well below the average values for the corresponding 65-74 age group (Activity and Health Research Limited, 1990 p. 81). The Allied Dunbar averages were approximately 3 W/kg for men and 1.9 for women, with 2 W/kg being the power necessary cited to climb stairs without assistance. Most trial participants provided values below this 2W/kg limit and the results reflect the difference in power and function between total knee joint arthroplasty patients and their non-arthroplasty peers.

Leg extensor press values following knee arthroplasty have been presented by Lamb and Frost (2003) (n=79). A preoperative mean for the operated side of 62 watts and 3/12 postoperative mean of 71 watts were obtained. These correspond with a preoperative mean of

46.6 watts and 3/12 mean of 69.5 watts for the current trial; the pre operative means were therefore lower whilst the 3/12 means were similar.

One limitation of the test was caused by the test procedure. The seat position was standardised for each patient. However, the post operative power values from patients with significant pre operative fixed flexion deformities (usually lessened during surgery) appeared adversely impacted upon by using this preoperative seat position.

The main disadvantage of using the leg extensor press however was its lack of portability. Of all the outcomes used in the trial the leg extensor press was the only one requiring hospital based follow up assessment. As already mentioned, the need to travel to hospital both prevented potential participants agreeing to participate in the trial and raised the risk of withdrawals from the trial. The need to offer flexible follow up arrangement, with home assessment where necessary, is considered more important to a future trial than the inclusion of the leg extensor press.

#### **4.8.4.4 Timed walk.**

The timed walk test was able to be completed by the majority of participants. Missing data was from participants unable to attend for follow up rather than from participants being unable to perform the test. Since the test is portable, it could be performed in any follow up assessment setting unlike the leg extensor press. The physiotherapy outcome assessors spoke of the seemingly high face validity of the test; participants appeared to strongly agree that measuring walking was meaningful.

It is difficult to find comparable trials using this outcome after knee joint arthroplasty. Frost *et al.*, (2002) used a 10 meter timed walk test (n=47) but timed people walking as quickly as possible. Their baseline walking speed of 1.18 m/s and 12/12 speed of 1.49 m/s was therefore expected to exceed the comfortable walking speeds measured in this trial. This

was a correct assumption, the current trial's walking speed was 0.71m/s at baseline and 1m/s at twelve months. Moffet *et al.*, (2004) used a six minute walking test which is discussed in the section comparing the current trial to previous trials below.

#### **4.8.4.5 Timed sit to stand.**

The value of the sit to stand test has been already highlighted in its introduction in the preceding chapter. It is fundamental to transfers and the initiation of gait (Schaubert and Bohannon, 2005) and correlates with leg press strength (Jones *et al.*, 1999), knee extensor strength, ambulatory independence, walking speed and stair climbing performance (Bohannon, 1998; Bohannon and Eriksrud, 2001; Ikezoe *et al.*, 1997).

The physiotherapy outcome assessors, like previous researchers, observed that knee arthroplasty patients find sit to stand a difficult manoeuvre and compensated, both pre and post operation, by taking more weight through the non affected side and by increasing forward leaning (Su *et al.*, 1998).

Again it is difficult to find comparable trials using this outcome after knee joint arthroplasty. The preoperative mean of 4.64 compared to a previous older community resident adults means of 12.7 (women) and 13.7 (men) again reflect the poorer functioning of this group compared to their non joint arthroplasty peers (Jones *et al.*, 1999).

By using the timed sit to stand version of the test, rather than the time taken to perform a specific number of repetitions, all participants who underwent objective testing (n=105) were able to provide test scores unlike 24% of knee osteoarthritis sufferers in previous research (Lin *et al.*, 2001).



#### **4.8.5 HOW DOES THIS TRIAL COMPARE TO PREVIOUS TRIALS?**

As stated in the systematic review of trials investigating the effectiveness of physiotherapy exercise following total knee joint arthroplasty for osteoarthritis, there are not many previously existing trials in this area.

This trial differs from previous ones in that it is the first trial to include both task specific training and functional exercises in the content of the intervention. Some previous trials have included descriptions of functional weight bearing exercises (Frost *et al.*, 2004; Kramer *et al.*, 2003; Mockford and Beverland, 2004; Moffet *et al.*, 2004). Whilst gait re-education and sit to stand are sometimes included in these descriptions the training in specific activities of daily living, such as car transfers and walking outside, have not been included previously. These activities were included since they had been identified as problematic activities for this patient group from published post operative Oxford knee scores (Dawson *et al.*, 1998).

In addition to the content being different, the dosage of the physiotherapy provided in the current trial is much briefer than in previous trials. Whilst the recommended repetition rate and frequency of exercises programmes were similar in this trial to previous trials, the physiotherapy – patient face to face contact time was far less. The data from the case report forms completed by the physiotherapists treatment patients provides medians of 10 and 11 minutes of estimated times spent with patients on functional exercises and task specific training respectively in visit one, and 8 and 13minutes in visit two. In the only previous trial demonstrating significant between group differences for function the dose was 12 outpatient physiotherapy sessions in 6-8 weeks; of 60-90 minutes duration per session (Moffet *et al.*, 2004). This amount of physiotherapy, given the numbers of knee arthroplasties occurring each year, would add up to be a significant proportion of out-patient physiotherapy time for the

NHS to fund. The current trial intervention, as a starting point, was deliberately designed to be a brief physiotherapy intervention which would be feasible for the NHS to consider implementing should it prove to be effective. No treatment differences have been found and, as mentioned, it may be that the dosage was insufficient and requires amendment.

The non significant findings in the current trial differ from the results of the systematic review of previous trials included earlier in this thesis and this discussion provides reasons for this finding. The earlier feasibility study by Frost *et al.*, (2002) suggested functional exercise may be superior to traditional exercise and the trial by Moffet *et al.*, (2004) showed significant differences in walking test distance, pain, stiffness and difficulties performing activities of daily living for their intervention group. In addition to the factors already discussed, a further difference between Moffet *et al.*, and the current trial may be the choice of tool/approach used to measure outcomes. For example, Moffet *et al.*, used a six minute walk test rather than a timed 10 m walk test as used in the current trial. The much lengthier test of endurance may have been more sensitive in picking up changes between participants than the short 10 m version. The objective 6 minute timed walk test was also the primary outcome in the Moffet *et al.*, trial, rather than the self report Oxford knee score used in the current trial, and perhaps the use of sufficiently sensitive objective functional measures may be preferable when measuring physiotherapy interventions over a general knee measure of orthopaedic surgery? It is difficult to speculate when the current trial was underpowered. As highlighted in the systematic review, the need for further high quality trials addressing the questions surrounding physiotherapy after knee joint arthroplasty appears evident.

#### **4.8.6 IMPLICATIONS FOR CLINICAL PRACTICE.**

The trial's findings showed no differences between the two types of physiotherapy provided in this trial so what does this mean for clinical practice? Although this early trial was never expected to provide definitive results regarding the type and dose of physiotherapy there are still some interesting considerations regarding clinical practice to consider.

The patients in this trial were believed to have no physiotherapy routinely organised following discharge, as per hospital protocol. Usual physiotherapy care was expected to mean that patients continued with the advice and exercises provided to them during their hospital stay. It was expected that a few percent of patients with complications might occasionally be treated /reviewed after discharge as appropriate. The results from the patient resource diaries however showed that 17-18% of patients in both groups received further physiotherapy. A conservative total of 212 visits were recorded; an average of over 4 per person receiving treatment. The size of this number of additional treatments, and the variety of sources from which additional rehabilitation was obtained, was a surprise. No readily available data could be found regarding the amount, sources or costs of physiotherapy currently provided to following discharge after knee joint arthroplasty in the UK. Whilst anecdotally it is known that some Trusts routinely provide 6 sessions of outpatient physiotherapy to patients following discharge and some purport to provide none routinely, the true amount of physiotherapy currently provided is not only unknown but impossible to guess. This research suggests that the costs, even for areas not routinely planning post discharge physiotherapy, may be being spread amongst many different sources and be much greater than currently believed. The true cost of rehabilitation to the NHS and to patients/employers funding rehabilitation is unknown. Whilst it is known that knee joint arthroplasty patients experience considerable functional impairment post operatively when compared with their peers (Noble

*et al.*, 2005) research is needed to establish both the current costs and the cost benefits of post knee joint arthroplasty rehabilitation to determine whether rehabilitations costs are worthwhile in terms of benefit to patients. Until it is known if, or how much, post discharge physiotherapy is required to benefit patients the NHS cannot decide whether this cost is acceptable and reasonable or not worth funding in the future.

With regard to the intervention itself, it is still believed that this is a potentially promising approach which needs further research before its value can be accepted or refuted. The approach was based upon a systematic review (Minns Lowe *et al.*, 2007) and was as evidence based as it was possible to make at the time. Whilst there were no differences observed between the two trial arms, the trial was impacted upon by some important factors which could potentially have masked any treatment trends occurring in the home visit group. By chance, the majority of serious co-morbidities developing during the trials affected the home visit group. These co-morbidities, for example hydrocephalus and rheumatoid arthritis were all of a type to severely impact upon participants' function and activities of daily living. Similarly, the breakdown of major complications occurring in both groups shows the home visit group to contain more complications likely to affect outcome (for example, both knees requiring revision surgery, more strokes) than the usual physiotherapy group (more chest infections and pressure sores). Also, although the trial excluded participants planning further joint arthroplasty during the trial period, twelve patients subsequently underwent such surgery; nine of whom were in the home visit group. Whether this was because participants in this group felt they recovered more quickly and were thus ready for further surgery, as some participants stated to the trial physiotherapists, was not investigated in this study. However, such surgery was also likely to impact upon the outcomes recorded for these participants at follow up assessments. When all these factors are considered together it may perhaps be

thought surprising that the home visit group did not fare worse than the usual physiotherapy group. The possibility that trends/treatment effects might have been present but were obscured by these differences occurring between the two groups has arisen but it is beyond the scope of the current work to explore this further at this time.

#### **4.8.7 FUTURE RESEARCH.**

It is planned to use the current trial data to update the systematic review in chapter three. It is also believed that the current trial has demonstrated the need for a future adequately powered trial investigating the role of physiotherapy upon functional outcome following discharge from hospital following physiotherapy. Any single trial is on a limited number of patients, is liable to encounter some methodological difficulties and usually requires replication (Pocock, 1983 p 93). This is certainly true for the current underpowered early trial. The current trial is considered to have served as a successful pilot trial to inform the development of such a future trial. It is believed to have done this in the following ways:

1. By providing information to enable adequate sample size to be determined.
2. By considering a more flexible approach to assessment follow up. By enabling follow ups to occur at patient's homes as well as hospitals, at the patients preference, recruitment and retention is likely to be optimised.
3. By refining the intervention. There are several options for a future trial. Firstly, to run a trial along similar lines as the current trial that is adequately powered. However, with concerns regarding the dosage of the intervention, it is worth considering amending the intervention so that a higher dose is received by participants. A further option, since the trial shows some patients do receive additional physiotherapy, would be to design a comparison of two active physiotherapy intervention groups, rather than

using a no routinely organised group. This would rather depend on the results of a possible survey (see below).

4. By refining the choice of outcomes. As mentioned it is believed that the alteration of the 10 m timed walking test to a six minute walking test may be indicated. Also, to remove the leg extensor press to remove the necessity of hospital based follow up. The use of hand held muscle dynamometry could be explored instead.

The current trial showed that a significant proportion of patients receive additional physiotherapy from a variety of sources when no physiotherapy is routinely organised. The research highlights that what physiotherapy patients receive following joint arthroplasty in the UK, and its content, is actually unknown. A patient survey to obtain information regarding the extent and nature of post discharge physiotherapy after knee joint arthroplasty would both provide information about current practice and provide an estimation of the level of resources currently consumed by this area of care. This information could also be used to develop a more comparable treatment arm against which to investigate a future functional exercise and task specific training physiotherapy trial.

The measurement of complications following knee joint arthroplasty has been indicated as another area that would benefit from further research. The development of measures and procedures to obtain accurate and relevant data regarding complications is apparent. The current uncertainty and lack of reporting regarding complications, and complication rates, that matter to patients and health care professionals cannot be allowed to continue for such a common orthopaedic procedure. Patients should be able to expect to be provided with clearer information regarding complications when deciding whether to undergo major surgery.

## **4.9 CHAPTER SUMMARY.**

This chapter explained the choice of methodological approach, the development of a trial intervention, and presented the trial protocol and results. The trial is an innovative trial to start to explore the role of functional exercise and task specific training in patients undergoing primary elective total knee arthroplasty for osteoarthritis.

The alternate hypotheses for this trial, that participants receiving two additional physiotherapy home visits after discharge from hospital following knee joint arthroplasty would show greater improvement in Oxford Knee Scores and secondary outcomes than participants receiving usual physiotherapy care, were rejected.

The trial null hypotheses, that no statistically significant difference in Oxford Knee and secondary outcome scores at 3,6, and 12 months post operation time points between participants receiving two physiotherapy home visits and participants receiving usual physiotherapy care would be observed, are accepted.

The trial proved to be underpowered with the presence of potential confounders affecting the home visit physiotherapy group more than the usual physiotherapy group. Therefore, although the lack of differences found between groups might still be due to both treatment arms providing similar outcomes, this cannot be assumed at this stage. A future, adequately powered trial, whose altered methodology allows for a higher dose of the intervention and the impact of future surgery, additional physiotherapy and the development of co-morbidities, is required before an accurate answer can be obtained.

# **5 CHAPTER FIVE. BLIND OUTCOME ASSESSMENT IN A PRAGMATIC PHYSIOTHERAPY REHABILITATION TRIAL.**

## **5.1 CHAPTER OVERVIEW.**

This chapter will commence by providing definitions of terms used with blinding in clinical trials and outlining the ambiguities within existing definitions. The background to the subject is then presented, followed by an overview of the effects on clinical trials of blinding and unblinding. The feasibility of achieving successful blind outcome assessment in trials and the difficulties this creates are then raised, and existing measurement and analysis approaches discussed. Following this, the purpose and methodology of a study exploring the feasibility of achieving blind outcome assessment in a pragmatic physiotherapy rehabilitation trial involving older people is detailed. Results are presented next, followed by a discussion of the study's findings.

## **5.2 INTRODUCTION.**

Whilst it was not possible to blind participants and clinicians providing the trial intervention regarding treatment allocation, it was possible to blind the outcome assessor. As will be shown later in this chapter, there is uncertainty regarding how best to measure blind outcome assessment with few trials providing blind outcome assessment success rates or adequate details regarding blind outcome assessment. This study provided the opportunity to enable detailed reporting of blind outcome assessment procedures and rates as recommended by the



CONSORT statement and also the opportunity to explore further the feasibility of achieving successful blind outcome assessment in pragmatic rehabilitation trials.

### **5.3 DEFINITIONS AND TERMINOLOGY.**

Blinding refers to the process of concealing group identity after treatment assignment through randomisation (Schulz *et al.*, 2002). Sometimes the term masked is used rather than blinding (Schulz *et al.*, 2002). Terminology surrounding blinding remains varied and confusing; a survey of physicians and textbooks found that for one widely used term, double blind, 17 interpretations were provided by physicians and 9 by textbooks (Devereaux *et al.*, 2001). Similar confusion exists for published trials; one review identified 8 varying definitions of a double blind trial amongst 83 trials (Montori *et al.*, 2002); another found 15 different operational meanings for the term amongst 200 trials (Haahr and Hróbjartsson, 2006).

Generally, trials are labelled open, single, double or triple blind trials.

Open trials are where both participant and investigator are aware of treatment allocation (Day & Altman, 2000). The term open trial may also be referred to as a non blind trial and may also denote a trial where the participant, investigator and outcome assessor are all aware of treatment allocation (Schulz *et al.*, 2002).

Single blind trials are where only either the participant or the investigator is blind to treatment allocation (Day & Altman, 2000; Forder *et al.*, 2005), or where only one of the participant, investigator or blind outcome assessor are aware (Schulz *et al.*, 2002).

Double blind trials are where both participants and investigator are blinded to treatment allocation (Day & Altman, 2000; Forder *et al.*, 2005) or where two of the participant, investigator or blind outcome assessor are aware (Schulz *et al.*, 2002) or sometimes all three (Fergusson *et al.*, 2004a).

Triple blind trials describe either a double blind trial plus blind data analysis, or a trial where participant, investigator and outcome assessor are all blinded (Schulz *et al.*, 2002). The latter being the same as Fergusson *et al's.*, (2004a) definition of a double blind trial. In addition, the investigator and the outcome assessor may be the same person so a triple blind trial might also be a double blind trial depending on which definition is used!

It is likely that the definitions by Schulz *et al.*, (2002) supersede those by Day and Altman (2000), since Altman co-authored both papers, but later definitions by Fergusson *et al.*, (2004a) and Forder *et al.*, (2005) illustrate continuing confusion. Schulz *et al.*, (2002) reflect that the “fuzziness reflects true ambiguity” and urges authors to explicitly state what blinding occurred in a trial rather than solely using single/double/triple blind terminology; a view concurred with by Devereaux *et al.*, (2001) and Haarh and Hróbjartsson, (2006).

#### **5.4 BACKGROUND.**

Blinding has been used in medical research for over 200 years with the intention of reducing bias and false treatment effect estimates, preventing participant attrition, and reducing the occurrence of co-interventions/additional therapeutic interventions within trials (Altman *et al.*, 2001; Schulz *et al.*, 2002). Bias in this chapter is used to describe systematic errors which encourage one outcome over any others and, in particular, blinding aims to minimise bias associated with researchers and participants expectations (Gluud, 2006). The use of blinding is generally well accepted (Davis Eyler *et al.*, 1999). However, frustratingly it is often erroneously assumed that a double blind trial must be of higher quality than a single blind trial even when double/triple blinding is impossible to achieve (Schulz *et al.*, 2002).

Although clear reporting of blinding evaluations was recommended in the original CONSORT statement to improve the reporting of clinical trials (Altman *et al.*, 2001) the vast

majority of published trials do not include evaluations of blinding (Sackett, 2007; Bang *et al.*, 2004) and even when included, details are often missing or poorly reported (Altman *et al.*, 2001; Montori *et al.*, 2002; Schulz and Grimes, 2002; Hróbjartsson *et al.*, 2007). In a review of 200 trials published in 2001, none of the trials claiming to be double blind (n=156) were considered to have completely reported the blinding status of key trial personnel and only 2% explicitly reported the blinding status of participants, providers and data collectors. (Haahr and Hróbjartsson, 2006). However it must not be assumed that trials which do not report blinding have not used blinding measures to reduce bias, even if this is the assumption within many systematic reviews and openly advised by some researchers (Schulz and Grimes, 2002): when contacted directly, many trial authors supply full and adequate details regarding blinding processes (Hill *et al.*, 2002; Devereaux *et al.*, 2004). Additionally it does not follow that the use of blinding processes includes measuring blinding success. When Hróbjartsson *et al.*, (2007) contacted the authors of 200 trials not containing details regarding blinding success only 12% (n=15) of 130 respondents replied they had formally measured blinding success. The same authors found that only 31/1599, or 2%, of blinded randomised clinical trials indexed in the Cochrane Central Register of Controlled Trials published in 2001 reported tests for the success of blinding. The authors discuss the possibility of the occurrence of selective reporting, with underreporting of unsuccessful blinding because of the impact poor blinding might have upon the perceived validity of the trial, in addition to lack of testing. The size of their study, plus its inclusion of trials both reporting and not reporting results, make this study one of the most useful ones to date in identifying the scale of the issues surrounding blinding in clinical trials.

The proportion of published trials in leading orthopaedic journals providing inadequate details regarding blinding has been reported as 54% (Bhandari *et al.*, 2002).

Poolman *et al.*, (2007) found that sixteen of the 32 orthopaedic trials included in their review (which contained one physiotherapy trial) did not report blinding of outcome assessors when blind outcome assessment was considered possible. This appears lower than the 72% of 50 non pharmacological trials reporting the occurrence of outcome assessment blinding when this was considered feasible by Boutron *et al.*, 2004. The difference may be due to the different study populations. This trial again found that blind outcome assessment is problematic for many rehabilitation and similar trials; for the subgroup of rehabilitation trials (n=27) the proportion of trials where blind outcome assessment was considered possible was only 22%.

A further issue is that it is also unclear what should be done if blinding processes prove inadequate since, although there is then potential for bias to occur, it does not follow that this is necessarily the case (Prescott *et al.*, 1999). Trials with lapses in blindness should not necessarily be dismissed as poor since this may be the case for many, if not most, trials because true blinding is so difficult to achieve (Turner *et al.*, 2002). In a random sample of 200 trials published in high repute journals, the success of blinding was described as less than optimal in nine of fourteen trials reporting information regarding blinding (Fergusson *et al.*, 2004). Even in the trials reporting blinding success this was considered debatable by the reviewers. Hróbjartsson *et al.*, (2007) also found that in 6 out of 7 trial reports they identified containing findings which indicated a loss of blinding, the risk of bias due to unblinding was either disregarded or ignored by authors. Clear interpretations regarding blinding related bias were evidently not occurring in these trial reports and appear to be generally omitted. This may be partly due to the difficulties researchers experience in assessing what determines successful blinding since it seems there is little mention of what

constitutes poor, acceptable or successful blinding in the literature which makes it hard for researchers to adequately measure and report findings.

#### **5.4.1 EFFECTS OF BLINDING.**

A discussion by Day and Altman (2000) explains that the relevance and appropriateness of blinding varies according to circumstances and the type of intervention under investigation. Generally trials using inadequate blinding show larger treatment effects than effectively blinded studies. Also, blind outcome assessment can often be more important than blind treatment administration, especially when using subjective outcomes, since lack of blind outcome assessment can result in systematic differences in outcome measurement known as ascertainment bias (Boutron *et al.*, 2006).

Studies vary in their findings regarding the impact of non blinded studies upon treatment effect. Some studies reviewing the effects of non blind outcome assessment have demonstrated significantly larger estimates of treatments ranging from ratio of odds ratios 0.31 (95% CI 0.2 to 0.47) (Poolman *et al.*, 2007). Research has also reported that general non blinding can exaggerate treatment effects by 20-45% relative to the true treatment effect which may be important for trials, such as physiotherapy trials, seeking small to moderate effects in trials (summarised by Devereaux, *et al.*, 2004). Other research reports lesser impact upon treatment efficacy. A previous meta-analysis, relating key aspects of trial quality to effect estimates by Jüni *et al.*, (2001a), estimated that open trials exaggerate treatment effect by 14%. This lower figure may in part be due to the inclusion of a study showing no effect from lack of double blinding. Additionally, an earlier meta-analysis of quality within trial reports found a small absolute difference of 3.7% scores higher for masked than unmasked

scores, compared to 30-40% exaggerated treatment efficacy for trials with inadequate allocation concealment (Moher *et al.*, 1998). These authors subsequently query whether the effort required to achieve successful double blinding in trials is worthwhile. The low quality of the trials included in this meta-analysis may have contributed to its findings. Therefore, while bias may often exaggerate treatment effect size this cannot be assumed; additionally the direction of bias may be difficult to ascertain since this is highly unpredictable (Fergusson *et al.*, 2004a; Gluud, 2006).

Schulz *et al.*, (2002) summarise the effects of bias produced by non/poor blinding as follows: knowledge of the intervention can affect treatment response, participants generally assume a new intervention is better than those already existing, adherence to new interventions may be greater than for established treatment/control arms, and the beliefs of investigators can influence participants and outcome assessment (a view concurred by Davis Eyler *et al.*, 1999). Many authors consider that subjective outcomes are considered more prone to this bias than objective ones (Day and Alman, 2000; Schulz *et al.*, 2002; Schultz and Grimes, 2002; Veira and Bangdiwala 2007). Hróbjartsson *et al.*, (2007) and Boutron *et al.*, (2006) add the possibilities of creating differing tendencies to drop out of the trial, in seeking out non-protocolised additional treatments, and in differing degrees of the attention provided to participants to this list of potential biases.

#### **5.4.2 FEASIBILITY OF ACHIEVING BLIND OUTCOME ASSESSMENT**

The feasibility and relevance of blinding varies between pharmaceutical and non pharmaceutical trial interventions (Boutron *et al.*, 2003; Day & Altman, 2000 ). Double blind trials, difficult to achieve even in pharmaceutical trials (Turner *et al.*, 2002), are often impossible for orthopaedic trials involving interventions such as surgery and physiotherapy

(Poolman *et al.*, 2007; Schulz and Grimes, 2002). In these trials it is frequently impossible to blind patients and care providers and care providers themselves are usually an integral part of the intervention (Boutron *et al.*, 2003). Such trials need to be judged on their merit rather than an inappropriate and inapplicable standard based upon the use of double blinding (Schulz and Grimes, 2002). Sham interventions or withholding treatment may also be unethical as well as impossible (Boutron *et al.*, 2003). Although double blinding may not be possible it is usually possible and necessary to use blind outcome assessment to prevent/limit biased outcome assessment (Poolman *et al.*, 2007; Prescott *et al.*, 1999; Forder *et al.*, 2005).

It is difficult to establish the feasibility of achieving successful blinding when, as already identified, so few trials report whether blind outcome assessment occurred let alone the blinding rates achieved. Boutron *et al.*, (2004) examined 50 non pharmacologic trials (including 27 rehabilitation and similar trials of which 14 were related to physiotherapy) and determined the blinding of participants to be feasible in 22% of the rehabilitation and similar non pharmacologic trials, the blinding of care providers to be feasible in 18% and the blinding of outcome assessors to be feasible in 22%. For the current trial participants and care providers are not blinded regarding treatment assignment. There is mixed feasibility regarding outcome assessment. Although the majority of secondary trial outcomes are objective measures, for which blind outcome assessment is feasible, the primary outcome is a self report measure and the patients are aware of their treatment allocation. Boutron *et al.*, (2003) state blind outcome assessment is impossible to achieve when patients are unblinded to treatment assignment and the primary outcome is a self report measure. Their associated belief that blind outcome assessment is achieved when a blind patient completes a self report outcome (Boutron *et al.*, 2004) has been refuted by Poolman *et al.*, (2007) who have

demonstrated that, unless the administrator of the self report measure is also blind, larger treatment effects occur.

It is difficult to assess blinding accurately. Attempts to measure blinding may be measuring treatment hunches; for example there may be an assumption that improving trial participants are receiving the active intervention (Sackett, 2004, 2007). Where two patients in different trial arms have equal responses to treatment the patient assumed to be in the active treatment group may be rated more highly (Carroll et al 1994). In this way outcome influences blinding responses. Studies report variety in their reported blind outcome results. Davis Eyer *et al.*, (1999) found that guesses from their blind evaluator team were correct only slightly more often than chance (56%). Carroll *et al.*, (1994) report their two outcome assessors to have correctly guessed treatment allocation with greater accuracy (76%). As Boutron *et al.*, (2005) discuss, some researchers consider blinding to be successful if the proportion of guesses is no better than chance. However an unequal proportion of guesses is also possible in well blinded trials; for example correct guesses may be higher when an intervention is proving successful so the usefulness of guessing is debatable. Measurement is therefore not straightforward.

#### **5.4.3 MEASUREMENT AND ANALYSIS OF BLIND OUTCOME ASSESSMENT**

There is no consensus regarding the optimum method to measure blindness in clinical trials since no validated measure has yet been developed and agreed. This creates considerable uncertainty and variety regarding how best to measure blinding and whom should be tested: participants or trial personnel (Hróbjartsson *et al.*, 2007; Boutron *et al.*, 2005). Although indices have been published (Bang *et al.*, 2004) they have not been widely accepted or reported in subsequent trial reports. These are not appropriate for use in the current trial since



they are designed to measure participant's beliefs regarding treatment assignment and the participants in this trial are aware of their assignment. Similarly, the minimum set of reported blinding information required for each trial suggested by Fergusson *et al.*, (2004a) also includes the counts of participants correctly guessing treatment assignment. Hróbartsson *et al.*, (2007) also suggest asking key trial persons to guess treatment allocation, enabling the comparison of this data to be compared with the actual treatments, to assess blinding success.

Fergusson *et al.*, (2004a) suggest that the minimum set of blinding methods and information, to be reported by trialists, consists of a reported count of participants guessing rates, inclusion of analytical methods and results used to evaluate the extent to which blinding is considered successful, plus an interpretation of the efficacy of blinding and any effects on study results. These suggestions have provoked many responses on the publishing journal's comments pages (11 to date) criticising the validity of counting guesses with some even advocating no during/after trial measurements of blinding. This lack of consensus regarding measurement creates difficulties for researchers who need to be able to report and interpret any blinding processes; to enable readers to determine the extent to which blinding related bias may have been introduced into studies.

Published blinding methods have been summarised by Boutron *et al.*, (2006, 2007). They describe the use of sham procedures, blinding participants to the trial hypothesis and the use of blinded centralization of the primary outcome. Sham exercise for the knee does not exist and was not considered possible for the current trial. It was considered that blinding participants to the trial hypothesis until the end of the trial would have been unethical for the current trial since patients needed to be aware of the time commitment required by the intervention programme prior to deciding whether to participate. It also would not have been possible to prevent contact between the two trial arms since patients may meet on the hospital

wards. The resources available for this trial prevented blinded centralization of the primary outcome via video, audiotape or photography or by someone not associated with the trial. The outcome assessor was employed to also approach potential participants so needed to know the purpose of the study.

Another approach is to develop a specific trial protocol designed to minimise possible causes of unblinding (Davis Eyler *et al.*, 1999). Carroll *et al.*, (1994) follow this approach recommending that outcome assessors are limited in the amount of trial information they are allowed to handle, participants are thoroughly informed about the importance of blinding and repeatedly requested not to reveal their treatment allocation, and safeguards are put in place to minimise unblinding. Additionally, procedures should be established at the beginning of the trial for dealing with any instances of unblinding (Forder *et al.*, 2005). Turner *et al.*, (2002) add to this list of recommendations by suggesting the study expectations and beliefs of involved personnel, such as outcome assessors, also needs to be explored.

An alternative to asking categorical Yes/No/Don't know responses to questions and guesses regarding blinding is to ask participants and/or outcome assessors and/or investigators to rate their certainty regarding their treatment assignment guess on a scale from 0-10 (Turner *et al.*, 2002). This option does not allow for "do not know" responses however and excludes data from people with this opinion. Since "do not know" rates may be high in a successfully blinded trial, (although this may not always be true) it can be argued that these responses, and also the rate of non response, do need to be reported and included in data analyses where possible (Boutron *et al.*, 2005). In addition, Boutron *et al.*, (2005) feel that it should be reported whether participants are forced to guess or allowed to express uncertainty regarding blinding to assist accurate interpretation of results. Rees *et al.*, (2005) have studied "do not know" responses at multiple time points during a trial investigating the efficacy of water

treatment units (179 trial participants). They obtained forced guesses from participants providing “do not know” responses and found that “do not know” respondents at the beginning of their trial held similar beliefs regarding group allocation to those providing unprompted responses, whereas forced guesses at the end of their trial were more consistent with random guesses.

The time points at which blinding should be tested is a further area of debate. Trials are inconsistent regarding when, and how often, blinding measurement occurs (Boutron *et al.*, 2005). Sackett (2004) and Walter *et al.*, (2005) advocate testing for blindness prior to trials, not during or after them, on the grounds that it is not possible to separate the efficacy of blinding from the effects of pre-trial hunches regarding efficacy. These authors however are referring to double blind pharmacological trials where it is possible to blind participants and clinicians. This is not possible for non pharmacologic trials and the majority of published reports for these trials involve measuring blinding during or at the end of trials (Bhandari *et al.*, 2002; Poolman *et al.*, 2007; Boutron *et al.*, 2004). In addition, measuring blindness prior to the trial does not permit the extent to which blinding was maintained through the trial to be evaluated (Hróbjartsson *et al.*, 2007). The difficulty of measuring blinding at the end of the study, however, is that assessment may be influenced by observations and assumptions regarding efficacy and symptoms which may cause bias (Walter *et al.*, 2005; Sackett, 2007). The presence of a true intervention effect would also impact upon measures of blindness at this time and, as already mentioned, a lack of blinding does not necessarily mean the bias has been introduced into the study (Hróbjartsson *et al.*, 2007). Some trials measure blindness only once whilst others measure it several times to provide information at multiple timepoints and allow for changes in blinding status to be picked up more easily (Boutron *et al.*, 2005). Measuring blindness solely at the end of a trial in order to identify significant group

differences due to unblinding depends upon the assumption that participants beliefs regarding treatment allocation are undecided at the start of the trial (Rees *et al.*, 2005). A recent study by Rees *et al.*, (2005) studied changes in participant's beliefs regarding treatment allocation at multiple time points during a trial, and found little difference between measuring twice, early and late trial, and measuring six times during the trial. They recommend testing at the two time points to prevent repeated questioning drawing attention to treatment allocation and risking increased unblinding, a risk in their own study. Their study also indicated that participants tended to initially believe they were assigned to the intervention group and that many participants altered their beliefs during the trial (31% or 41/132). They argue that bias may not only occur when participants correctly guess allocation but also when intergroup differences regarding treatment group allocation differentially effect outcome.

In their review identifying blinding methods used in randomised clinical trials, Boutron *et al.*, (2005) found that the most frequently reported method of analysis was counts of participant's guesses (79% or n=43). Statistical analyses were included in 57% (n=31) trial reports, usually to compare the proportion of correct guesses among those produced by chance in each group or globally or to use a chi-square test to check for a relation between participants' guesses and group allocation. Trials including uncertainty, for example by allowing participants to state "do not know" did not usually analyse this data statistically. While the authors promoted the use of indices to allow this data to be statistically analysed they found the use of such indices were rarely reported and, for this trial, no index for use with blind outcome assessors rather than participants could be found. In their review of blinding in 94 psychiatric trials Fergusson *et al.*, (2004) identified few reports giving any details regarding statistical data analysis; chi square tests (n=3), kappa values (n=2) and Fishers exact test (n=1) results were provided in six studies.

#### **5.4.4 SECTION SUMMARY.**

The use and value of blinding within randomised clinical trials is generally, but not universally, accepted. Ambiguities in defining terms used to describe blinding exist and it is not clear what constitutes successful blinding in reported trials. Uncertainty regarding the effects and feasibility of blinding and regarding how and when to best measure and analyse blinding data are evident.

#### **5.5 PURPOSE OF STUDY.**

The purpose of this study was to explore the feasibility of achieving blind outcome assessment in a pragmatic physiotherapy rehabilitation trial involving older people and to contribute to the limited available knowledge in this area. This was achieved by the following two objectives.

1. To record and present the number of instances of unblinding occurring during the trial.
2. To fully document details surrounding each instance of unblinding and to present a content analysis of the results.

#### **5.6 METHOD.**

##### **5.6.1 SECTION OVERVIEW.**

This section provides a description of the methods used to minimise instances of unblinding and to record all instances of unblinding occurring in the trial.

### **5.6.2 Protocol to minimise instances of unblinding.**

Since the blind outcome assessor was involved in approaching participants she had to be aware of the trial question and design. She reported no strong preference for either trial arm and no strong beliefs or expectations about the outcome of the trial. She did not see all participants prior to randomisation/intervention and blinding was not measured at this time. It was possible to measure outcome assessment blinding at the three and twelve month follow up time points and therefore this approach was chosen.

All trial participants were informed about blind outcome assessment and its importance during their baseline assessment. All participants were requested not to reveal their group allocation on the following occasions:

1. Baseline assessment; both immediately prior to, and after, randomisation.
2. During the two home visits provided to the Home Visit treatment group.
3. On each occasion when participants were telephoned to arrange follow up visits.
4. At the commencement of each follow up visit.

In addition the following safeguards were followed. The outcome assessor had no access to any study data which could compromise blind outcome assessment. They had no access to the password protected study database or to completed assessments, treatment notes or questionnaires which were secured in a locked filing cabinet. Although it is recommended that blind outcome assessors should not come into contact with those colleagues providing the intervention (Boutron *et al.*, 2007) this was not possible since staff shared the same office and space restrictions prevented further office space from being available. For this reason three other safeguards were put in place:

1. The names and locations of home visits were not mentioned in the office.

2. Telephone calls to arrange visits were not made when the outcome assessor was in the office.
3. Incoming telephone calls to the office were screened. Either other staff answered the telephone or, if this was not possible, the outcome assessor would determine which of the Unit's trials the caller was telephoning about and arrange for the trial co-ordinator to telephone the participant back for callers involved in the present trial.

### 5.6.3 Recording instances of unblinding.

Every follow up assessment questionnaire contained the following question.

This question asks you whether you are aware/ unaware of which treatment group this patient belongs to in the trial.

*(Please tick one box only)*

- I do not know which group the patient is in.....
- I have guessed the patient is in the home treatment group .....
- I have guessed the patient is in the usual care group .....
- The patient has told me they are in the home treatment group.....
- The patient has told me they are in the usual care group.....

This question included both guesses and knowledge because the impact of guessing allocation is likely to be different than that of actually knowing the intervention (Boutron *et al.*, 2004).

Unless the outcome assessor stated they did not know to which group the participant had been allocated they recorded why/how they had guessed/knew group allocation in a field

diary kept for the purpose. Each instance was recorded on a separate page and sealed to ensure the outcome assessor could not revisit previous instances to lessen the chance of the assessor remembering allocation when next seeing the participant.

## **5.7 DATA ANALYSIS.**

The questionnaire data was analysed descriptively and the results presented using bar charts and tables. As previously mentioned there are no existing indices or accepted other approaches available for use to evaluate the opinions of blind outcome assessors. The diary data was fully transcribed and categories identified using content analysis and constant comparison; the data was read and re-read to identify these categories following which each item was checked and compared against all others to establish the categories as truly reflective of the data (Pope *et al.*, 2000). The findings from this process were then summarised in a table.

## **5.8 RESULTS.**

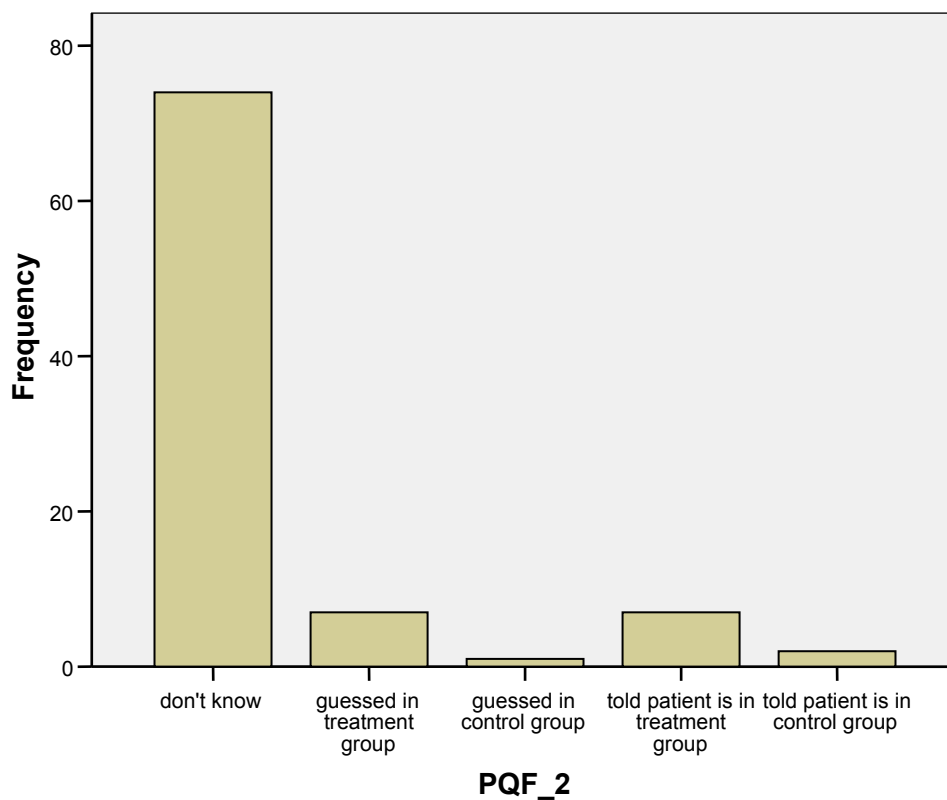
The vast majority of all outcome assessments at all time points were performed by the same assessor. A few assessments only were performed by another blinded and experienced research physiotherapist in the rare event of the lead assessor being unavailable.

### **5.8.1 Questionnaire Findings.**

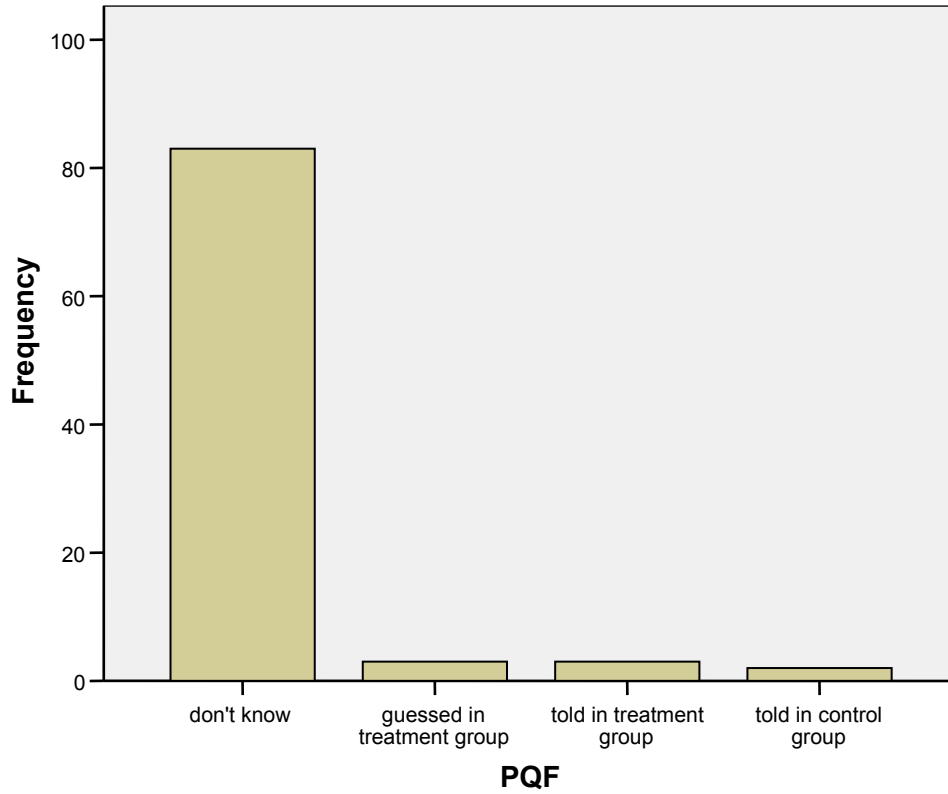
The bar charts below provide a visual representation of blind outcome assessor beliefs regarding treatment arm allocation at three and twelve month face-to-face follow up assessments. Since randomisation occurred after baseline assessment and six month follow up



was a postal questionnaire there are no data for these time points. The tables following this show the frequencies of correct and incorrect answers. These tables show successful blind outcome assessment occurring for the majority of assessments: 74/91 (81.32%) of participants at 3/12 and 83/91 (91.21%) at 12/12 follow ups therefore fewer instances of unblinding occur at the 12/12 time point. At both timepoints a few incorrect guesses/telling occur: 3 instances at 3/12 and 2 at 12/12.



**Figure 27. Bar chart to show the outcome assessor's beliefs regarding treatment allocation at three month follow up assessments.**



**Figure 28.** Bar chart to show the outcome assessor's beliefs regarding treatment allocation at twelve month follow up assessments.

**Table 25.** Table to show the numbers of correct and incorrect guesses of the outcome assessor's beliefs regarding treatment allocation at three month follow up assessments.

		Percent	Frequency	Frequency of correct guesses	Frequency of incorrect guesses
Valid	don't know	74	69.2		
	guessed in treatment group	7	6.5	6	1
	guessed in control group	1	.9	1	
	told patient is in treatment group	7	6.5	6	1
	told patient is in control group	2	1.9	1	1
	Total	91	85.0		
Missing	999	2	1.9		
	Number not assessed	14	13.1		
	Total	16	15.0		
Total		100.0	107		

**Table 25. Table to show the numbers of correct and incorrect guesses of the outcome assessor’s beliefs regarding treatment allocation at twelve month follow up assessments.**

		Percent	Frequency	Frequency of correct guesses	Frequency of incorrect guesses
Valid	don't know	83	77.6		
	guessed in treatment group	3	2.8	2	1
	told in treatment group	3	2.8	2	1
	told in control group	2	1.9	2	
	Total	91	85.0		
Missing	999	3	2.8		
	Number not assessed	13	12.1		
	Total	16	15.0		
Total		100.0	107		

### **5.8.2 Diary Data.**

40 instances (28 participants) of unblinding were recorded in the Unblinding Events Diary during the trial by the outcome assessor. The table below summarises the results of the content analysis of the diary data and presents the main reasons/causes of unblinding incidents. Not all unblinding instances provided correct information: 12.5% of events lead to the outcome assessor reaching the wrong conclusion regarding to which treatment arm an individual had been allocated. In addition, it was also mentioned that three participants revealed their allocation close to the end of their assessment and one after the assessment has finished but before the participant had left.

**Table 26. Correct and incorrect instances of unblinding documented in assessor's diary (40 instances in 28 subjects)**

Methods of unblinding	Correct instances of unblinding: assessor informed/guessed correctly	Incorrect instances of unblinding: assessor informed/guessed incorrectly
Revealed by patient:		
Patient telling assessor	18	
Patient asking additional queries		1
Patient mentioning other physiotherapy /follow ups		3
Patient phoning to rearrange appointment times	2	
Patient leaving randomisation card outside envelope	1	
Revealed by relative/friend:		
Relative/friend telling assessor	3	1
Revealed by colleagues:		
During conversation with physiotherapy colleagues	4	
Overhearing office conversations /phone calls	3	
Assessor receiving phone call about a specific participant's needs	1	
Assessor answering query from another member of staff	1	
Seeing trial paperwork:		
Home diary from patient	1	
Patient summary with randomisation	1	
<i>Total</i>	35	5

## **5.9 DISCUSSION.**

### **5.9.1 SECTION OVERVIEW.**

This section contains the discussion of the methodology and findings of the study evaluating and exploring the feasibility and success of achieving blind outcome assessment in a pragmatic physiotherapy rehabilitation trial. The section begins with a summary of the key findings of the study. It then provides a discussion of the strengths and limitations of the study's procedures followed by a discussion of the implications of the study and directions for future research.

### **5.9.2 SUMMARY OF KEY FINDINGS.**

The study shows that successful blind outcome assessment rates were obtained during this pragmatic physiotherapy rehabilitation trial. Successful blinding rates of over 80% of assessments at three month follow up, and over 90% at twelve month follow up, were achieved. The main cause of unblinding events was participants telling the assessor. The findings show that not all unblinding events are remembered at subsequent assessments, even in this relatively small trial.

### **5.9.3 STRENGTHS AND LIMITATIONS OF THE STUDY'S PROCEDURES.**

As highlighted in the background section in chapter four, the vast majority of published trials either do not include evaluations of blinding (Sackett, 2007; Bang *et al.*, 2004) or evaluations are poorly reported (Altman *et al.*, 2001; Montori *et al.*, 2002; Schulz and Grimes, 2002; Hróbjartsson *et al.*, 2007). This study therefore is one of the few available

musculoskeletal rehabilitation studies providing data regarding blind outcome assessment for a randomised clinical trial. By providing outcome assessment blinding rates, in addition to explaining why the participants and clinicians providing the trial intervention cannot be blinded, this study goes beyond the newly extended CONSORT statement detailing standards for the reporting of pragmatic trials (Zwarenstein *et al.*, 2008).

The questionnaires findings contained 81-91% “do not know” responses from the outcome assessor. This high value is believed to reflect the success of using a specific trial protocol to minimise unblinding events as used successfully by Carroll *et al.*, (1994) and as recommended by Davis Eyler *et al.*, (1999).

As explained in the trial protocol, it was not possible to blind trial participants and physiotherapists to treatment assignation. It was noted on the trial database that several usual physiotherapy care trial arm participants vocalised disappointment with their group assignation at the time of randomisation, stating a preference for the home visit package. The presence of any systematic biases resulting from the strength of patient preferences or beliefs about the treatment efficacy of the two physiotherapy trial packages is unknown. A patient preference trial, where participants with a strong preference to a specific trial arm are assigned to that trial arm while people without preferences are randomised as usual, is one option to avoid this form of bias (Halpern, 2003). This option was considered during the design stage of the trial but not incorporated because since such trials are only partially randomised and therefore the trial’s main results could be affected by uncontrolled confounding (Halpern, 2003).

In line with Davis Eyler *et al.*, (1999) our study also had instances where comments were made that instantly unblinded the outcome assessor. Unlike the previous study however, only one blind outcome assessor was available for each follow up session so it was not

possible for another blind outcome assessor to take the place of the unblinded one. It was felt unreasonable to ask participants to re-attend at another time since, unlike Davis Eyer *et al.*, (1999) reassessments could not be performed at patients homes. Whilst the majority of participants telling their assessor their group allocation did so accurately, one interesting point arising from the study was that two participants at 3/12 and one participant at 12/12 directly told the assessor their group allocation but provided erroneous information. This may have been due to a mishearing / misunderstanding on the part of the assessor, confusion on the part of the participant or a direct attempt to confuse/mislead assessor by participants. It does mean however that not all statements from patients regarding group allocation are true and assessors may believe they have been unblinded when true group allocation has not been revealed.

Concern has been expressed regarding the use of repeated measures in studies since it is feared that repeatedly asking for beliefs regarding blinding may draw attention to the topic and make people consider, remember or analyse the issue more often (Rees *et al.*, 2005; Boutron *et al.*, 2004). This study not only used repeated measures but asked for detailed description of all instances of unblinding. We believed it was improbable for repeated measures to effect our non blinded participants but there was a concern that the assessor might remember instances of unblinding from one time point to another; especially since this trial is relatively small. Although this may have occurred to some extent the substantially higher twelve month successful blinding rates than those at three months suggest that fully recording unblinding events during the trial did not seriously and adversely effect blinding rates at final follow up. This adds useful information to the current debate regarding when and how often blinding measurements should occur (Boutron *et al.*, 2005).

The variation of the timing of the majority of unblinding incidents is interesting. There were more instances of unblinding occurring earlier in the trial and at the three month

assessment time point than later on and at twelve months. There are several possible explanations for this. Whilst some participants repeatedly divulged their group allocation, it is possible that some participants did subsequently learn not to mention their group allocation. In addition, the diary data does identify instances of unblinding occurring because outcome assessment and clinical staff had to share an office. Boutron *et al.*, (2007) recommends that blind outcome staff should not come into contact with clinicians to prevent such instances but initially this was not possible since additional office space was unavailable. However, the main outcome assessor changed jobs partway through the trial. This meant that, after completing the vast majority of the 3/12 assessments and only a few 12/12 assessments, she moved to a different hospital site only returning to carry out assessments. Since only one instance of unblinding was revealed by colleagues after this job move took place this would appear to emphasize the importance of keeping the outcome assessor away from the research office and clinical environment where possible.

The diary data usefully supplemented the questionnaire data in a further way. Although the questionnaire question included both guesses and being told group allocation as Boutron *et al.*, (2004) recommends, the question used only allowed assessors to state whether unblinding occurred. It did not mention when the unblinding incident occurred. The diary data recorded several instances when participants revealed their allocation near the end of the assessment, with one participant revealing their allocation after all the assessment measures had been completed. The question used in the questionnaire therefore would not, in its current form, distinguish between participants informing assessors when such information is likely to influence results from situations when little impact is likely to have been caused.



#### **5.9.4 IMPLICATIONS OF THE STUDY.**

It is difficult to place the results from this study into context with results from other physiotherapeutic or orthopaedic trials due to the lack of previously published trial blinding rates. However Boutron *et al.*, (2005) helpfully state that a high number of “do not know” answers is to be expected in a successfully blinded trial. The 81-91% “do not know” responses obtained in this study are considered high and the trial to have successfully achieved blind outcome assessment.

The combination of using a blind outcome question in follow up questionnaires together with a diary to record instances of unblinding has not been found elsewhere. We believe it is new to this study. The use of this combination has revealed interesting findings. The diary data includes the instances of unblinding not captured via the assessment questionnaires (for example, overhearing colleagues discussing the trial or during telephone calls with patients). It suggests that instances of unblinding occurring in the office environment or revealed by colleagues were less remembered than patients directly informing the assessor during assessments. It has also identified a limitation of the questionnaire question, the issue of distinguishing between participants informing assessors when such information is likely to influence results from times when little impact is likely to have been caused. The completion of an unblinding diary adds little to the workload of the outcome assessor. The length of entries suggests that all the useful information can be obtained via 2-3 minutes of writing time. The use of a diary allows any type of event to be recorded spontaneously and expressively, whilst avoiding the limitations of closed questions or preordained categories (Oppenheim, 1992 p.113-114).

Since undertaking this study a provocative commentary has been published which questions the whole validity and possibility of measuring blindness; arguing that blinding

should be relegated to the level of process and the emphasis shifted to testing for bias-generating consequences for any loss of blinding instead (Sackett, 2007). This identifies several bias generating consequences which apply to the current study. Firstly, the level of contamination caused by participants in the usual physiotherapy trial arm obtaining the intervention elsewhere since they know they will not receive it from the trial. Since participants were requested to detail any additional physiotherapy they received external to the trial in their patient diaries it has been possible to assess the extent to which this occurred and this is discussed in chapter four. Secondly, if physiotherapists who are aware of treatment allocation provide effective treatments to the usual physiotherapy trial arm participants this again could cause bias. The design of the current trial, which assessed the efficacy of an extra intervention in addition to any usual physiotherapy received, made this concern unlikely. In addition the physiotherapists providing the trial intervention did not treat any of the participants in any other setting. Thirdly, if participants or physiotherapists know or have hunches regarding their trial arm they may downplay or under report symptoms/adverse outcomes. Although this can not be ruled out for the current trial, the detailed information participants provided regarding symptoms and complications makes this appear unlikely.

This commentary by Sackett (2007) may reflect an important thought shift in the research community. The CONSORT statement has been revised for pragmatic trials to remove the sentence regarding the necessity of evaluating and reporting blindness; this has been toned down to providing an explanation of the approach used (Zwarenstein *et al.*, 2008). It is presently unclear whether this change is being driven by increasing awareness that researchers are simply unwilling to measure/report blindness rates or by an increasing acceptance that it is pointless to do so. The recent publication of further papers calling for detailed reporting regarding blindness, such as Poolman *et al.*, (2007) mean this issue may become a contentious topic of trial design in the next few years. Schulz (2002) opines that “investigators must not only minimise bias but must also communicate those efforts to the reader”. Hróbjartsson *et al.*, (2007) state that the critical appraisal of blinding in most trials is meaningless because of grossly incomplete reporting. The long awaited revised CONSORT statement for pragmatic trials (Zwarenstein *et al.*, 2008) seems disappointingly vague, missing the opportunity to facilitate true development in the measurement and reporting of blinding in pragmatic trials. In the meantime, reporting full information regarding blinding methods and results as suggested by Fergusson *et al.*, (2004a) seems a sensible approach. Particularly if, as also suggested, trialists provide readers with an interpretation of the results. Such interpretations might prevent possible errors in understanding, such as those recounted by Sackett (2007), from occurring. As Fergusson *et al.*, (2004 b) argue, although the methods by which blinding is currently measured could do with improvement, to subsequently decide to ignore blinding and not even attempt to assess it is a poor decision. Clinicians seeking evidence based information upon which to make treatment decisions are still often forced to base their decisions on trials with unclear control of bias (Gluud, 2006). This study allows us to

provide researchers and clinicians with results regarding the effectiveness of blinding procedures used in the physiotherapy trial earlier in this thesis. These results, plus their interpretation, should enable readers to assess for themselves the extent to which blinding has been successful and bias minimised.

#### **5.9.5 FUTURE DIRECTIONS.**

From a physiotherapy perspective the paradigm shift raised by Sackett (2007) would appear potentially positive. If trial evaluations and systematic review scoring systems were subsequently revised to allow the extent to which bias generating concerns from loss of blinding to be assessed in trials, instead of immediately downgrading the quality of a trial because it is not possible for double blinding to occur, then this could promote improved quality assessments in future physiotherapy trials. Blind outcome assessment, plus additional forms of blinding when possible and appropriate, could remain integral to trial design plus increased explicit consideration would then occur regarding possible consequences of loss of blinding.

Patient preferences regarding treatment allocation may have influenced this trial. Halpern (2006) suggests that one approach to include patients' preferences for treatment arms, whilst avoiding the pitfalls of a patient preference trial design, is to assess participants' preference prior to subsequently randomising them in the normally accepted way. Preference can then at least be included and quantified in the trial's statistical analyses. Although not yet common within physiotherapy/therapeutic trials this is not unknown (for example Klaber-Moffett *et al.*, 1999, Kalauokalani *et al.*, 2001) and results have appeared mixed. Preference affected outcome in one of these trials (Kalauokalani *et al.*, 2001) but not the other (Klaber-Moffett *et al.*,

1999). Further research is required to determine whether the benefits of quantifying preference bias in this way are worth the additional data collection and analysis.

As already mentioned resource restrictions prevented the use of centralised outcome assessment advocated by Boutron *et al.*, (2005). However this is a possibility to consider in future trials. If this approach was included it would also then permit the exploration of the impact of blinding in the way that Noseworthy *et al.*, (1994) performed. By asking both blinded and non blinded outcome assessors to assess efficacy Noseworthy *et al.*, (2004) were able to demonstrate the impact of blinding in their trial; physician blinding prevented an erroneous conclusion regarding treatment efficacy in their trial and it would be very interesting to investigate the impact of outcome assessor blinding within future physiotherapy trials in this way.

Hróbjartsson *et al.*, (2007) stress the need for both improved methodologies and improved reporting of blinding by clinical trialists. It is hoped that this study may assist the development of an improved methodology. It is likely that, after future research, that the categories developed from this, and future studies, diary data could be used to develop a coding system for responses which would then enable responses to be quantified (Oppenheim, 1992 p.113-114). This, together with the questionnaire question used in this study, plus a further question to identify whether unblinding occurred before, during or following assessment might form the basis of a quick and easy to complete tool by which to measure blind outcome assessment in pragmatic trials. Since the majority of existing blinding measures relate to the blinding of participants such a tool would appear to be a valuable addition to available methodologies.

## **5.10 Chapter Summary.**

After presenting definitions and background research, this chapter presented the protocol, results and discussion of a study exploring the feasibility of achieving successful blind outcome assessment in a rehabilitation trial involving older people. The study found that blind outcome assessment was feasible and that a high rate of blinding was achieved.

## **6 CHAPTER SIX. A SUMMARY OF THE MAIN FINDINGS, IMPLICATIONS AND FUTURE RESEARCH SUGGESTIONS IN THIS THESIS.**

### **6.1 CHAPTER OVERVIEW.**

This chapter will summarise the main findings for the systematic review, the randomised clinical trial and the study evaluating blind outcome assessment. For each of these three component pieces of research the main implications for clinicians and researchers, and the priorities for future research will also be summarised. Finally, the conclusions of this thesis will be presented.

### **6.2 SYSTEMATIC REVIEW.**

#### **6.2.1 Summary of main findings.**

The main purpose of this review, to evaluate the effectiveness of post discharge physiotherapy exercise on function, walking, range of motion, quality of life and muscle strength, for osteoarthritic patients following elective primary total knee arthroplasty, was successfully achieved.

Six trials were identified, five of which were suitable for inclusion in meta-analyses. A 0.33 (95% CI 0.07 to 0.58) small to moderate standardised effect size, in favour of functional exercise, was seen for function at 3-4 months post operatively. Small to moderate weighted mean differences of 2.9 (95% CI 0.61 to 5.2) for range of joint motion and 1.66 (95% CI -1 to 4.3) for quality of life were seen, in favour of

functional exercise, at 3-4 months post operatively. Post-treatment benefits were not carried through to one year.

### **6.2.2 Implications of the Research.**

It would appear reasonable to refer patients for a short course of physiotherapy following discharge to provide short term patient benefit. The small to moderate standardised effect size obtained for function, favouring the intervention, is considered clinically important. This reflects actual improvements in one or more important aspects of function reported by patients after receiving the treatment intervention. The type of physiotherapy provided to patients following discharge after knee replacement surgery needs consideration. Traditional exercise programmes may be less effective in the short term following total knee arthroplasty than functional physiotherapy exercise interventions.

### **6.2.3 Future Research Suggestions.**

Where an effect is shown it is not huge indicating that although there are not many studies and they are not large they are still likely to have detected most worthwhile effects. These tentative findings suggest that further research would be worthwhile to reduce the level of uncertainty currently present and thus contribute to the acknowledged knowledge gap in post knee replacement rehabilitation. The content, dosage and timing of physiotherapy rehabilitation interventions following knee joint arthroplasty for osteoarthritis all warrant further investigation.



### **6.3 RANDOMISED CLINICAL TRIAL.**

#### **6.3.1 Summary of main findings.**

The alternate hypotheses for this trial, that participants receiving two additional physiotherapy home visits after discharge from hospital following knee joint arthroplasty would show greater improvement in Oxford Knee Scores and secondary outcomes than participants receiving usual physiotherapy care, were rejected.

The trial null hypotheses, that no statistically significant difference in Oxford Knee and secondary outcome scores at 3,6, and 12 months post operation time points between participants receiving two physiotherapy home visits and participants receiving usual physiotherapy care would be observed, are accepted.

The trial proved to be underpowered with the presence of potential confounders affecting the home visit physiotherapy group more than the usual physiotherapy group.

#### **6.3.2 Implications of the Research.**

This trial is an innovative, early, trial to start to explore the role of functional exercise and task specific training in patients undergoing primary elective total knee arthroplasty for osteoarthritis. The intervention itself is still believed to be a potentially promising approach which needs further research before its value can be accepted or refuted. The approach was based upon a systematic review (Minns Lowe *et al.*, 2007) and was as evidence based as it was possible to make at the time. Whilst there were no differences observed between the two trial arms, the trial was impacted

upon by some important factors which could potentially have masked any treatment trends occurring in the home visit group.

In addition, the current trial showed that a significant proportion of patients receive additional physiotherapy from a variety of sources when no physiotherapy is routinely organised. The research highlights that what physiotherapy patients receive following joint arthroplasty in the UK, and its content, is actually unknown.

### **6.3.3 Future Research Suggestions.**

There were three main areas for future research. Firstly it is believed that the current trial has demonstrated the need for a future adequately powered trial investigating the role of physiotherapy upon functional outcome following discharge from hospital following physiotherapy. This trial has also informed the development of such a future trial.

Secondly, a patient survey to obtain information regarding the extent and nature of post discharge physiotherapy after knee joint arthroplasty would both provide information about current practice and provide an estimation of the level of resources currently consumed by this area of care.

Thirdly, the development of measures and procedures to obtain accurate and relevant data regarding complications following knee joint arthroplasty has been indicated as another area that would benefit from further research. The current uncertainty and lack of reporting regarding complications, and complication rates, that matter to patients and health care professionals requires to be addressed.

## **6.4 BLIND OUTCOME ASSESSMENT IN A PRAGMATIC PHYSIOTHERAPY REHABILITATION TRIAL.**

### **6.4.1 Summary of main findings.**

This study successfully explored the feasibility of achieving blind outcome assessment in a pragmatic physiotherapy rehabilitation trial involving older people. Successful blinding rates of over 80% of assessments at three month follow up, and over 90% at twelve month follow up, were achieved. The main cause of unblinding events was participants telling the assessor. The findings show that not all unblinding events are remembered at subsequent assessments, even in this relatively small trial

### **6.4.2 Implications of the Research.**

This study is one of the few available musculoskeletal rehabilitation studies providing data regarding blind outcome assessment for a randomised clinical trial and, by so doing, has to contribute to the limited available knowledge in this area.

The study also contributes useful information to the current debate regarding when and how often blinding measurements should occur. The substantially higher twelve month successful blinding rates than those at three months suggest that the use of repeated measures, plus fully recording unblinding events during the trial, did not seriously and adversely effect blinding rates at final follow up.

This study used the combination of using a blind outcome question in follow up questionnaires together with a diary to record instances of unblinding. This dual approach has not been found elsewhere but, with each methodology providing both different and complementary information has proved a valuable methodology to achieve successful measurement. The questionnaire allowed blinding rates to be

established: the diary data captured incidents of unblinding the questionnaire was unable to record and identified that the questionnaire question was limited in its ability to distinguish between participants informing assessors when such information is likely to influence results from times when little impact is likely to have been caused.

### **6.4.3 Future Research Suggestions.**

Existing research demonstrates the need for new and improved methodologies to measure blinding in trials. The majority of available measures are for use when measuring the extent to which trial participants remain blind during trials with scant attention paid to the measurement of blind outcome assessment. It is hoped that similar diary data from future studies could be used to develop a coding system for responses which would then enable responses to be quantified (Oppenheim, 1992 p.113-114). This, together with the questionnaire question used in this study, plus a further question to identify whether unblinding occurred before, during or following assessment might form the basis of a quick and easy to complete tool by which to measure blind outcome assessment in pragmatic trials. It is believed this would be of value to other researchers undertaking pragmatic trials in addition to facilitating the reporting of blind outcome assessment rates for the benefit of those reading and applying future trials to their clinical practice.

## **6.5 CONCLUSIONS OF THIS THESIS.**

### **6.5.1 SYSTEMATIC REVIEW.**

The results of the systematic review to evaluate the effectiveness of physiotherapy exercise following knee arthroplasty for osteoarthritis are not conclusive; they indicate however that functional exercise interventions following discharge show promise for this patient group. Effect sizes were small to moderate. No long term benefit was found. Further research in this area therefore would appear justifiable and worthwhile.

From the evidence to date, it is proposed that functional exercises should be considered for inclusion in post discharge physiotherapy programmes following knee arthroplasty.

### **6.5.2 THE RANDOMISED CLINICAL TRIAL.**

No significant statistical differences were observed between the home visit physiotherapy and usual physiotherapy groups for the primary outcome, the Oxford Knee Score, and all secondary outcomes at three, six and twelve months post discharge from hospital following knee joint arthroplasty.

This early trial was underpowered. Whilst there were no differences observed between the two trial arms, the trial was impacted upon by some important factors which could potentially have masked any treatment trends occurring in the home visit group.

A future, adequately powered trial, using the findings from this current trial, is indicated to be worthwhile for this large patient group undergoing this common orthopaedic condition.

### **6.5.3 BLIND OUTCOME ASSESSMENT.**

Successful blind outcome assessment is both achievable and measurable in pragmatic rehabilitation trials. The use of repeated measures did not prove to be of concern in this instance. The use of specific trial protocols designed to minimise incidents of unblinding, together with a dual measurement approach of questionnaire questions plus a diary to fully detail every events of unblinding was used to achieve successful measurement of blind outcome assessment in this trial successfully in the trial in this thesis.

## **6.6 OVERALL THESIS.**

This thesis produced the first systematic review to evaluate the effective of physiotherapy exercise following joint replacement. It included the development of a new trial physiotherapy intervention, plus an early trial exploring this new rehabilitation approach. It also provided a new approach to the measurement of blind outcome assessment which has the potential for future development into a new research tool. Overall, this thesis has successfully contributed new knowledge to the area of physiotherapy rehabilitation following knee joint arthroplasty for patients with osteoarthritis.

## REFERENCES.

Abel, U., Koch, A. (1999) The role of randomization in clinical studies: myths and beliefs. *Journal of Clinical Epidemiology*, **52**(6):487-497.

Activity and Health Research Limited, *Allied Dunbar National Fitness Survey, 1990* [computer file]. Colchester, Essex: UK Data Archive [distributor], May 1995. SN: 3303.

Affleck, G., Tennen, H., Keefe, F.J., Lefebvre, J.C., Kashikar-Zuck, S., Wright, K., Starr, K., Caldwell, D.S. (1999) Everyday life with osteoarthritis or rheumatoid arthritis: independent effects of disease and gender on daily pain, mood and coping. *Pain*, **83**:601-609.

Altman, D. G. (1991) *Practical statistics for medical research*. London: Chapman & Hall.

Altman, D. G., Schulz, K. F., Moher, D., Egger, M., Davidoff, F., Elbourne, D., Gøtzsche, P. C., Lang, T. (2001) The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Annals of Internal Medicine*, **134**:663-694.

American Academy of Orthopaedic Surgeons guidelines  
<http://orthoinfo.aaos.org/topic.cfm?topic=A00357> [accessed 18/05/09].

Bang, H., Ni, L., Davis, C.E. (2004) Assessment of blinding in clinical trials. *Controlled Clinical Trials*, **25**(2):143-156.

Barker, K. L., Barrington, S. E., Clarkson-Webb, A. G., Squires, S., Racey, A. (2003) Pre-admission functional testing to predict length of stay in hospital after hip arthroplasty in a diagnostic and treatment centre setting. *Journal of Bone and Joint Surgery*, **90-B**, Issue SUPP\_I:11.

Barker, K., Lamb, S.E., Toye, F., Jackson, S., Barrington, S. (2004) Association between radiographic joint space narrowing, function, pain and muscle power in severe osteoarthritis of the knee. *Clinical Rehabilitation*, **18**:793-800.

Bassey, E. J., Fiatarone, M. A., O'Neill, E. F., Kelly, M., Evans, W. J., Lipsitz, L. A. (1992) Leg extensor power and functional performance in very old men and women. *Clinical Science*, **82**: 321-327.

Bassey, E. J., Short, A. H. (1990) A new method for measuring power output in a single leg extension: feasibility, reliability and validity. *European Journal of Applied Physiology*, **60**: 385-390.

Bassey, E. J., Tay, G., West, F. A. (1990) A comparison between measuring power output in a single leg extension and in weight-bearing activities of brief duration such as stair-running in man. *Journal of Physiology*, (London) **427**, 12P.

- Bassilico, F.C., Sweeney, G., Losina, E., Gaydos, J., Skoniecki, D., Wright, E. A., Katz, J. N. (2008) Risk factors for cardiovascular complications following total joint replacement. *Arthritis and Rheumatism*, **58** (7):1915-1920.
- Beaupré, L. A., Davies, D. M., Jones, C. A., Cinats, J. G. (2001) Exercise combined with continuous passive motion or slider board therapy compared with exercise only: a randomised controlled trial of patients following total knee arthroplasty. *Physical Therapy*, **81** (4):1029-1037.
- Beeston, S., Minns Lowe, C., Rastall, M. J. (2001) *“Mind the Gap!” An exploration of evidence – based practice among allied health professionals in three London Trusts*. University of East London: London.
- Bellamy, N., Buchanan, W. W., Goldsmith, C. H., Campbell, J., Stitt, L. W. (1988) Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *The Journal of Rheumatology*, **15** (12):1833-1840.
- Benedetti, M. G., Catani, F., Bilotta, T. W., Marcacci, M., Mariani, E., Giannini, S. (2003) Muscle activation pattern and gait biomechanics after total knee replacement. *Clinical Biomechanics*, **18**:871-876.
- Bennell, K. L., Hunt, M. A., Wrigley, T. V., Boon-Whatt, L., Hinman, R. S. (2008) Role of muscle in the genesis and management of knee osteoarthritis. *Rheumatic Disease Clinics of North America*, **34**:731-754.
- Berlin, J. A. (1997) Does blinding of readers affect the results of meta-analyses? *The Lancet* **350**:185-186.
- Bhandari, M., Richards, R.R., Sprague, S., Schemitsch, E.H. (2002) The quality and reporting of randomized trials in the journal of bone and joint surgery from 1998 through 2000. *Journal of Bone and Joint Surgery*, **84**(3):388-396.
- Black, N. (1996) Why we need observational studies to evaluate the effectiveness of health care. *British Medical Journal*, **312**:1215-1218.
- Bohannon, R. W. (1995) Sit-to-stand test for measuring performance of lower extremity muscles. *Percept Mot Skills*, **80**:163-166.
- Bohannon, R. W. (1998) Alternatives for measuring knee extension strength of the elderly at home. *Clinical Rehabilitation*, **12**: 434-440.
- Bohannon, R. W. and Eriksrud, O. (2001) What measures of lower extremity muscle strength best explains walking independence? *Journal of Physical Therapy Science*, **13**:1-3.
- Bolognesi, M. P., Marchant, M. H., Viens, N. A., Cook, C., Pietroban, R., Vail, T. P. (2008) The impact of diabetes on perioperative patient outcomes after total hip and knee arthroplasty in the Unites States. *The Journal of Arthroplasty*, **23** (6):92-98.



Bornman, J. (2004) The World Health Organisation's terminology and classification: application to severe disability. *Disability and Rehabilitation*, **26** (3):182-188.

Boutron, I., Guittet, L., Estellat, C., Moher, D., Hróbjartsson, A., Ravaud, P. (2007) Reporting methods of blinding in randomized trials assessing nonpharmacological treatments. *Public Library of Science Medicine*, **4**(2):0370-0380.

Boutron, I., Estellat, C., Ravaud, P. (2005) A review of blinding in randomized controlled trials found results inconsistent and questionable. *Journal of Clinical Epidemiology*, **58**(12):1220-1226.

Boutron, I., Estellat, C., Guittet, L., Dechartres, A., Sackett, D.L., Hróbjartsson, A., Ravaud, P. (2006) Methods of blinding in reports of randomized controlled trials assessing pharmacologic treatments: a systematic review. *Public Library of Science Medicine*, **3**(10):1931-1939.

Boutron, I., Tubach, F., Giraudeau, B., Ravaud, P. (2003) Methodological differences in clinical trials evaluating nonpharmacological and pharmacological treatments of hip and knee osteoarthritis. *Journal of the American Medical Association*, **290**(8):1062-1070.

Boutron, I., Tubach, F., Giraudeau, B., Ravaud, P. (2004) Blinding was judged more difficult to achieve and maintain in nonpharmacologic than pharmacologic trials. *Journal of Clinical Epidemiology*, **57**(6): 543-550.

Bowker, P., Messenger, N. (1988). The measurement of gait. *Clinical Rehabilitation*, **2**:89-97.

Bradley, F., Wiles, R., Kinmouth, A-L., Mant, D., Gantley, M. (1999) Development and evaluation of complex interventions in health services research: case study of the Southampton heart integrated care project (SHIP). *BMJ*, **318**:711-715.

Brander, V. A., Mullarkey, C. F. (2002) Rehabilitation after total hip replacement for osteoarthritis. *Physical Medicine and Rehabilitation: State of the Art Reviews*, **16**(3): 415-430.

Brandt, K. D., Radin, E. L., Dieppe, P. A., van de Putte, L. (2006) Yet more evidence that osteoarthritis is not a cartilage disease. *Annals of Rheumatic Diseases*, **65**:1261-1264.

Brandt, K. D., Dieppe, P., Radin, E.L. (2008) Etiopathogenesis of osteoarthritis. *Rheumatic Disease Clinics of North America*, **34**:531-449.

Brewin, C. R., Bradley, C. (1989) Patient preferences and randomised clinical trials. *British Medical Journal*, **299**:313-315.

British Orthopaedic Association and the Association for Surgery of the Knee (1999) *Knee replacement: a guide to good practice*. British Orthopaedic Association.

- Brosseau, L., Balmer, S., Tousignant, M., O'Sullivan, J. P., Goudreault, C., Goudreault, M., Gringras, S. (2001) Intra- and intertester reliability and criterion validity of the parallelogram and universal goniometers for measuring maximum active knee flexion and extension of patients with knee restrictions. *Archives of Physical Medicine and Rehabilitation*, **82**:396-401.
- Bulthuis, Y., Drossaers-Bakker, K. W., Taal, E., Oostenn, J., van't Pad Bosch, P., Oosterveld, F., van de Laar, M. (2007) Arthritis patients show long-term benefits from 3 weeks intensive exercise training directly following hospital discharge. *Rheumatology*, **46**:1712-1717.
- Callahan, C. M., Drake, B.G., Heck, D.A., Dittus, R. S. (1994) Patient outcomes following tricompartmental total knee replacement. A meta-analysis. *The Journal of the American Medical Association*, **271** (17):1349-1357.
- Campbell, M. J., Machin, D. (1993) *Medical statistics a commonsense approach*. John Wiley & Sons Ltd: Chichester.
- Carr, J. H. (1992) Balancing the centre of body mass during standing up. *Physiotherapy Theory and Practice*. **8**:159-164.
- Carroll, K.M., Rounsaville, B.J., Nich, C. (1994) Blind man's bluff: effectiveness and significance of psychotherapy and pharmacotherapy blinding procedures in a clinical trial. *Journal of Consulting and Clinical Psychology*, **62**(2):276-280.
- Chalmers, I (1993) The Cochrane Collaboration: preparing, maintaining and dissemination systematic reviews of the effects of health care. *Annals of the New York Academy of Sciences*, **703**:156-163.
- Chambers, I. (2001) Foreword p. xiii-xviii.  
In  
*Systematic reviews in health care Meta-analysis in context* (2001) Ed Egger, M., Davey Smith, G., Altman, D. BMJ Publishing Group: London.
- Codine, Ph., Delleme, Y., Denis-Laroque, F., Herisson, Ch. (2004) The use of low velocity submaximal eccentric contractions of the hamstring for recovery of full extension after total knee replacement: a randomised controlled study. *Isokinetics and Exercise Science*, **12**: 215-218.
- Cohen, J. (1960) A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, **20**:37-46.
- Cohen J. (1969) *Statistical power analysis for the behavioural sciences*. Academic Press: New York. p.23.
- Coutts, F.(2005) Total joint replacements. Chapter 8 p. 253-262.  
In  
*Physiotherapy in orthopaedics a problem solving approach Second Edition* (2005) Ed Atkinson, K., Coutts, F., Hassenkamp, A-M. Churchill Livingstone: London.

- Cowies, J., Khan, L. A., Ballantynes, J. A., Brenkel, I. J. (2009) The effect of smoking on the outcome after total hip replacement. *Journal of Bone and Joint Surgery*, **91-B**, Issue SUPP\_1,134.
- Creamer, P., Lethbridge-Cejku, M., Hochberg, M. C. (2000) Factors associated with functional impairment in symptomatic knee osteoarthritis. *Rheumatology*, **39**:490-496.
- Critical Appraisal Skills Programme. 10 questions to help you make sense of randomised clinical trials. <http://www.phru.nhs.uk/casp/rct.pdf> [accessed 14/05/05].
- Croft, P. (1998) Admissible evidence. *Annals of the Rheumatic Diseases*, **57**:387-389.
- Crombie, I. K., McQuay, H. J. (1998) The systematic review: a good guide rather than a guarantee. *Pain*, **76**: 1-2.
- D'Ambrosia, R. D. (2005) Epidemiology of Osteoarthritis. *Orthopedics*, **28** (2 Suppl):S201-205.
- Dauty, M., Genty, M., Ribinik, P. (2007) Physical training in rehabilitation programs before and after total hip and knee arthroplasty. *Annales de réadaptation et de médecine physique*, doi:10.1016/j.aarmp.2007.04.011.
- Davies, H. T. O., Crombie, I. K. (2001) What is meta-analysis?. [www.evidence-based-medicine.co.uk](http://www.evidence-based-medicine.co.uk), **1**(8):1-8.
- Davis Eyler, F., Behnke, M., Stewart, N., Wobie, K. (1999) Examiner masking in research of prenatal cocaine use. *Journal of Drug Issues*, **29** (2):215-224.
- Dawson, J., Fitzpatrick, R., Murray, D., Carr, A. (1998) Questionnaire on the perceptions of patients about total knee replacement. *Journal of Bone and Joint Surgery* **80**:63-9.
- Day, S.J., Altman, D.G. (2000) Statistics notes: blinding in clinical trials and other studies. *British Medical Journal* **321**(7259):504.
- Deeks, J. J., Altman, D. G., Bradburn, M. J. (2001) Statistical methods for examining heterogeneity and combining results from several studies in meta-analysis. Chapter 15 p. 285-312.  
In  
*Systematic reviews in health care Meta-analysis in context* (2001) Ed Egger, M., Davey Smith, G., Altman, D. BMJ Publishing Group: London.
- Deeks, J. J., Higgins, J. P. T., Altman, D. G. (2006). Analysing and presenting results.  
In  
Higgins, J. P. T., Green, S., editors. *Cochrane Handbook for Systematic Reviews of Interventions* 4.2.6 [updated September 2006]; Section 8. In: The Cochrane Library, Issue 4, 2006. John Wiley & Sons, Ltd: Chichester, UK.
- Dequeker, J., Luyten, F. P (2008) The history of osteoarthritis-osteoarthrosis. *Annals of the Rheumatic Diseases*, **67**:5-10.

Devereaux, P.J., Choi, P.T.L., El-Dika, S., Bhandari, M., Montori, V.M., Schünemann, H.J., Garg, A.X., Busse, J.W., Heels-Ansdell, D., Ghali, W.A., Manns, B.J., Guyatt, G.H. (2004) An observational study found that authors of randomized controlled trials frequently use concealment of randomization and blinding, despite the failure to report these methods. *Journal of Clinical Epidemiology*, **57**(12):1232-1236.

Devereaux, P.J., Manns, B.J., Ghali, W.A., Quan, H., Lacchetti, C., Montori, V.M., Bandari, M., Guyatt, G.H. (2001) Physician interpretations of textbook definitions of blinding terminology in randomized controlled trials. *The Journal of the American Medical Association*, **285**(15):2000-2003.

Dickersin, K., Scherer, R., Lefebvre, C. (1994) Identifying relevant studies for systematic reviews. *British Medical Journal* **309**:1286.

Dieppe, P., Basler, H-D., Chard, J., Croft, P., Dixon, J., Hurley, M., Lohmander, S., Raspe, H. (1999) Knee replacement surgery for osteoarthritis: effectiveness, practice variations, indications, and possible determinants of utilization. *Rheumatology*, **38**:73-83.

Dieppe, P., Judge, A., Williams, S., Ikwueke, I., Guenther, K-P., Floeren, M., Huber, J., Ingvarsson, T., Learmonth, I., Lohmander, L.S., Nilsson, N., Puhl, W., Rowley, D., Thieler, R., Dreinhoefer, K and EUROHIP study group. Variations in the pre-operative status of patients coming to primary hip replacement for osteoarthritis in European orthopaedic centres. *BMC Musculoskeletal Disorders*, **10**:19.

Dieppe, P. A., Lohmander, L. S. (2005) Pathogenesis and management of pain in osteoarthritis. *The Lancet*, **365**:965-973.

Dinah, A. F., Mears, S. C. (2008) Primary total knee arthroplasty in the elderly. *Current Opinion in Orthopaedics*, **19**:63-67.

Dixon, T., Shaw, M., Ebrahim, S., Dieppe, P. (2004) Trends in hip and knee joint replacement: socioeconomic inequalities and projections of need. *Annals of the Rheumatic Diseases*, **63**:825-830.

Dunbar, M. J., Robertsson, O., Ryd, L., Lidgren, L. (2001) Appropriate questionnaires for knee arthroplasty. Results of a survey of 3600 patients from the Swedish knee arthroplasty registry. *Journal of Bone and Joint Surgery*, **83-B**:339-344.

Edwards, C. (2002) A proposal that patients be considered honorary members of the healthcare team. *Journal of Clinical Nursing*, **11**:340-348.

Egger, M., Davey Smith, G. (1998) Meta-analysis bias in location and selection of studies. *British Medical Journal*, **316**:61-66.

Egger, M., Davey Smith, G. (2001b) Principles of and procedures for systematic reviews. Chapter 2 p. 23-42.

In

*Systematic reviews in health care Meta-analysis in context* (2001) Ed Egger, M., Davey Smith, G., Altman, D. BMJ Publishing Group: London.

Egger, M., Davey Smith, G., O'Rourke, K. (2001a) Rationale, potentials and promise of systematic reviews. Chapter 1. p. 3-19.

In

*Systematic reviews in health care Meta-analysis in context* (2001) Ed Egger, M., Davey Smith, G., Altman, D. BMJ Publishing Group: London.

Egger, M., Dickersin, K., Davey Smith, G. (2001c) Problems and limitation in conducting systematic reviews p. 43-68.

In

*Systematic reviews in health care Meta-analysis in context* (2001) Ed Egger, M., Davey Smith, G., Altman, D. BMJ Publishing Group: London.

Egger, M., Jüni, P., Bartlett, C., Holenstein, F., Sterne, J. (2003) How important are comprehensive literature searches and the assessment of trial quality in systematic reviews? Empirical study. *Health Technology Assessment*, **7**(1):1-82.

Enloe, L. J., Shields, R. K., Smith, K., Leo, K., Miller, B. (1996) Total hip and knee replacement treatment programs: a report using consensus. *Journal of Orthopaedic and Sports Physical Therapy*, **23**(1):3-11.

eorthopod.com/images.

[http://www.eorthopod.com/images/ContentImages/knee/knee\\_osteoarthritis/knee\\_osteoarth\\_intro01.jpg](http://www.eorthopod.com/images/ContentImages/knee/knee_osteoarthritis/knee_osteoarth_intro01.jpg) [accessed 18/05/09].

Eriksrud, O., Bohannon, R. W. (2003) Relationship of knee extension force to independence in sit-to-stand performance in patients receiving acute rehabilitation. *Physical Therapy*, **83**:544-551.

Eysenck, H. J. (1994) Systematic reviews: Meta-analysis and its problems. *British Medical Journal*, **309**:789-792.

Fautrel, B., Hilliquin, P., Rozenberg, S., Allaeart, F-A., Coste, P., Leclerc, A., Rossignol, M. (2005) Impact of osteoarthritis: results of a nationwide survey of 10,000 patients consulting for OA. *Joint Bone Spine*, **72**:235-240.

Fehring, T. K., Odum, S. M., Griffin, W. L., Mason, J. B., McCoy, T. H. (2007) The obesity epidemic its effect on total joint arthroplasty. *The Journal of Arthroplasty*, **22** (6) Suppl. 2:71-76.

Feinstein, A. R., Cicchetti, D. V. (1990) High agreement but low kappa. I. The problems of two paradoxes. *Journal of Clinical Epidemiology*, **43** (6):543-549.

Felson, D. T., Lawrence, R. C., Dieppe, P.A., Hirsch, R., Helmick, C. G., Jordan, J. M., Kington, R. S., Lane, N. E., Nevitt, M. C., Zhang, Y., Sowers, M., McAlindon, T., Spector, T. D., Poole, A. R., Yanovski, S. Z., Ateshian, G., Sharma, L., Buckwalter, J. A., Brandt, K.D., Fries, J. F., (2000a) Osteoarthritis: New insights Part 1: The disease and its risk factors. *Annals of Internal Medicine*, **133**:635-646.

- Felson, D. T., Lawrence, R. C., Dieppe, P.A., Hirsch, R., Helmick, C. G., Jordan, J. M., Kington, R. S., Lane, N. E., Nevitt, M. C., Zhang, Y., Sowers, M., McAlindon, T., Spector, T. D., Poole, A. R., Yanovski, S. Z., Ateshian, G., Sharma, L., Buckwalter, J. A., Brandt, K.D., Fries, J. F., (2000b) Osteoarthritis: New insights Part 2: Treatment approaches. *Annals of Internal Medicine*, **133**:726-737.
- Felson, D.F., Naimark, A., Anderson, J., Kazis, L., Castelli, W., Meenan, R. (1987) The prevalence of knee osteoarthritis in the elderly. The Framingham osteoarthritis study. *Arthritis & Rheumatism*, **30**(8):914-918.
- Felson, D. F., Neogi, T. (2004) Osteoarthritis: is it a disease of cartilage or of bone? *Arthritis & Rheumatism*, **50**:341-344.
- Fergusson, D., Cranley Glass, K., Waring, D., Shapiro, S. (2004a) Turning a blind eye: the success of blinding reported in a random sample of randomised, placebo controlled trials. *British Medical Journal*, **328**(7437):432-43.
- Fergusson, D., Cranley Glass, K., Waring, D., Shapiro, S. (2004b) Turning a blind eye authors' reply. *British Medical Journal*, **328**:1136.
- Field, A. (2000) *Discovering statistics using SPSS for windows*. SAGE Publications Ltd: London.
- Forder, P.M., Gebiski, V.J., Keech, A.C. (2005) Allocation concealment and blinding: when ignorance is bliss. *Medical Journal of Australia*, **182**(2):87-89.
- Friedman, L. M., Furberg, C. D., DeMets, D. L. (1998) *Fundamentals of Clinical Trials*. Springer-Verlag New York, Inc: New York.
- Frost, H., Lamb, S. E., Robertson, S. (2002) A randomised controlled trial of exercise to improve mobility and function after elective knee arthroplasty. Feasibility, results and methodological difficulties. *Clinical Rehabilitation*, **16**:200-209.
- Gaine, W. J., Ramamohan, N.A., Hussein, N. A., Hullin, M. G., McCreath, S. W. (2000) Wound infection in hip and knee arthroplasty. *The Journal of Bone and Joint Surgery*, **82-Br**:561-565.
- Garratt, A. M., Brealey, S., Gillespie, W. J., in collaboration with the DAMASK trial team (2004) Patient-assessed health instruments for the knee: a structured review. *Rheumatology*, **43** (11):1414-1423.
- Gluud, L.L. (2006) Bias in clinical intervention research. *American Journal of Epidemiology*, **163**(6):493-501.
- Gooberman-Hill, R., Woolhead, D., MacKichan, F., Ayis, A., Williams, S., Dieppe, P. (2007) Assessing chronic joint pain: lessons from a focus group study. *Arthritis and Rheumatism*, **57**(4):666-671.

Gøtzsche, P. C. (2000) Why we need a broad perspective analysis on meta-analysis. *British Medical Journal*, **321**:585-586.

Greenhalgh, T. (1997) Papers that summarise other papers (systematic reviews and meta-analyses) *British Medical Journal*, **315**:672-675.

Grégoire G., Derderian, F., Le Lorier, J. (1995) Selecting the language of the publications included in a meta-analysis: is there a tower of babel bias? *Journal of Clinical Epidemiology*, **48**(1):159-163.

Groll, D.L., To, T., Bombardier, C., Wright, J. G. (2005) The development of a comorbidity index with physical function as the outcome. *Journal of Clinical Epidemiology*, **58**:595-602.

Gursen, M. D., Ahrens, J. (2003) The key to a new home care protocol: prospective visits. *Caring Magazine*, **23**(1):40-44.

Guyatt, G.H., Pugsley, S.O., Sullivan, M. J., Thompson, P. J., Berman, L., Jones, N. L., Fallen, E. L., Taylor, D. W. (1984) Effect of encouragement on walking test performance. *Thorax*, **39**(11):818-822.

Haarr, M.T., Hróbjartsson, A. (2006) Who is blinded in randomized clinical trials?: a study of 200 trials and survey of authors. *Clinical Trials* **3**(8):360-365.

Halpern, S. D. (2003). Evaluating preference effects in partially unblinded, randomised trials. *Journal of Clinical Epidemiology*, **56**:109-115.

Hannan, M.T., Felson, D.T., Rincus, T. (2000) Analysis of the discordance between radiographic changes and knee pain in osteoarthritis of the knee. *The Journal of Rheumatology*, **27**:1513-1517.

Harcourt, W. G. V., White, S. H., Jones, P. (2001) Specificity of the Oxford knee status questionnaire. The effect of disease of the hip or lumbar spine on patients' perception of knee disability. *Journal of Bone and Joint Surgery*, **83-B**:345-47.

Hawker, G.G., Davis, A. M., French, M. R., Cibere, J., Jordan, J.M., March, L., Suarez-Almazor, M., Katz, J.N., Dieppe, P. (2008) Development and preliminary psychometric testing of a new OA pain measure – an OARSI/OMERACT initiative. *Osteoarthritis and Cartilage*, **16** (4):409-414.

Heck, D. A., Robinson, R. L., Partridge, C. M., Lubitz, R. M., Freund, D. A. (1998) Patient outcomes after knee replacement, *Clinical Orthopaedics and Related Research*, **356**:93-110.

Hewitt, B., Shakespeare, D. (2001) Flexion vs. extension: a comparison of post-operative total knee arthroplasty mobilisation regimes. *The Knee*, **8**:305-309.

Higgins, J. P. T., Green, S., Editors (2005a). *Cochrane Handbook for Systematic Reviews of Interventions* 4.2.5 [updated May 2005]. In: The Cochrane Library, Issue 3, 2005. John Wiley & Sons, Ltd: Chichester, UK.

- Higgins, J. P. T., Green, S. Editors (2005b) *Locating and selecting studies*. Cochrane Handbook for Systematic Reviews of Intervention 4.2.5 [updated May 2005]; Section 5. in: The Cochrane Library, Issue 3, 2005. John Wiley & Sons: Chichester, UK.
- Higgins, J. P. T., Thompson, S. G., Deeks, J. J., Altman, D.G.(2003) Measuring inconsistency in meta-analyses. *British Medical Journal*, **327**:557-560.
- Hill, C.L., LaValley, M.P., Felson, D.T. (2002) Discrepancy between published report and actual conduct of randomized clinical trials. *Journal of Clinical Epidemiology*, **55**(8):783-786.
- Hopayian, K. (2001) The need for caution in interpreting high quality systematic reviews. *British Medical Journal*, **323**:681-684.
- Hospital episode statistics online. Main operation summary table.  
<http://www.hesonline.nhs.uk/Ease/servlet/ContentServer?siteID=1937&categoryID=204> [accessed 6<sup>th</sup> May 2009].
- Howe, T., Oldham, J. (1995) Functional tests in elderly osteoarthritic subjects: variability of performance. *Nursing Standard*, **9**:35-8.
- Howell, D. C. (2002) Intraclass Correlation: For Unordered Pairs.  
[http://www.uvm.edu/~dhowell/StatPages/More\\_Stuff/icc/icc.html](http://www.uvm.edu/~dhowell/StatPages/More_Stuff/icc/icc.html) [accessed 18<sup>th</sup> May 2009].
- Hróbjartsson, A., Forfang, E., Haahr, M.T., Als-Nielsen, B., Brorson, S. (2007) Blinded trials taken to the test: an analysis of randomized clinical trials that report tests for the success of blinding. *International Journal of Epidemiology*, **36**(3):654-663.
- Hughes, K., Kuffner, L., Dean, B. (1993) Effect of weekend physical therapy treatment on postoperative length of stay following total hip and knee arthroplasty. *Physiotherapy Canada*, **45** (4):245-249.
- Hunter, D. J., Felson, D. T. (2006) Osteoarthritis. *BMJ*, **332**:639-642.
- Hunter, D. J., McDougall, J. J., Keefe, F.J. (2008) The symptoms of osteoarthritis and the genesis of pain. *Rheumatic Disease Clinics of North America*, **34**:623-643.
- Husted, H., Holm, G. (2006) fast tract in total hip and knee arthroplasty – experiences from Hvidovre university hospital, Denmark. *International Journal of the Care of the Injured*, **37S**:S31-S35.
- Ikezoe, T., Asakawa, Y., Hazaki, K., Kuroki, H., Morinnaga, T., Kawano, I., Kanzaki, H., Aoki, N. (1997) Muscel strength and muscle endurance required for independent walking in the elderly. *Journal of Physical Therapy Science*, **9**:19-22.
- Ioannidis, J. P.A., Haidich, A-B., Lau, J. (2001) Any casualities in the clash of randomised and observational evidence? *British Medical Journal*, **322**:879-880.



- Iyengar, K.P., Nadkarni, J.B., Ivanovic, N., Mahale, A. (2007) Targeted early rehabilitation at home after total hip and knee joint replacement: does it work? *Disability and Rehabilitation*, **29**(6):495-502.
- Jämsen, E., Huotari, K., Huhtala, H., Nevalainen, J., Konttinen, Y. T. (2009) Low rate of infected knee replacements in as nationwide series – is it an underestimate? *Acta Orthopaedica*, **80**:205-212.
- Jansen, Y.J.F.M., Bal, R., Bruijnzeels, M., Foets, M., Frenken, R., de Bont, A. (2006). Coping with methodological dilemmas; about establishing the effectiveness of interventions in routine medical practice. *BMC Health Services Research*, **6**:160.
- Janssen, W. G. M., Bussman, H. B. J., Stam, H. J. (2002) Determinants of the sit-to-stand movement: a review. *Physical Therapy*, **82**(9):866-879.
- Jette, A. M. (2006) Toward a common language for health, disability, and health. *Physical Therapy*, **86** (5):726-734.
- Jones, B., Jarvis, P., Lewis, J.A., Ebbutt, A. F. (1996) Trials to assess equivalence: the importance of rigorous methods. *British Medical Journal*, **313**:36-39.
- Jones, C. A., Beaupre, L. A., Johnstone, D. W. C., Suarez-Almazor, M. E. (2007) Total joint arthroplasties: current concepts of patient outcomes after surgery. *Rheumatic Disease Clinics of North America*, **33**:71-86.
- Jones, C. J., Rikli, R. E., Beam, W. C. (1999) A 30-s chair-stand test as a measure of lower body strength in community-residing older adults. *Research Quarterly for Exercise and Sport*, **70** (2):113-119.
- Jones, D.L., Westby, M.D., Greidanus, N., Johanson, N.A., Krebs, D.E., Robbins, L., Rooks, D.S., Brander, V. (2005) Update on hip and knee arthroplasty: current state of evidence. *Arthritis & Rheumatism*, **53**(5):772-780.
- Jones, J., Hunter, D. (1995) Qualitative research: consensus methods for medical and health services research. *British Medical Journal*, **311**:376-380.
- Jordan, K. M., Arden, N. K., Doherty, M., Bannwarth, B., Bijlsma, J. W. J., Dieppe, P., Gunther, K., Hauselmann, H., Herrero-Beaumont, G., Kaklamanis, P., Lohmander, S., Leeb, B., Lequesne, M., Mazieres, B., Martin-Mola, E., Pavelka, K., Pendleton, A., Punzi, L., Serni, U., Swoboda, B., Verbruggen, G., Zimmerman-Gorska, I., Dougados, M. (2003) EULAR recommendations 2003: an evidence based approach to the management of knee osteoarthritis: report of a task force of the standing committee for international clinical studies including therapeutic trials (ESCISIT). *Annals of the Rheumatic Diseases*, **62**:1145-1155.
- Jüni, P., Altman, D. G., Egger, M. (2001a) Assessing the quality of randomised clinical trials p. 87-108.  
In

*Systematic reviews in health care Meta-analysis in context* (2001) Ed Egger, M., Davey Smith, G., Altman, D. BMJ Publishing Group: London.

Jüni, P., Altman, D. G., Egger, M., (2001b) Assessing the quality of controlled clinical trials. *BMJ*, **323**:42-6.

Jüni, P., Dieppe, P., Donovan, J., Peters, T., Eachus, J., Pearson, N. *et al.*, (2003) Population requirement for primary knee replacement surgery: a cross sectional study. *Rheumatology*, **42**:516-521.

Jüni, P., Holenstein, F., Sterne, J., Bartlett, C., Egger, M. (2002) Direction and impact of language bias in meta-analyses of controlled trials: empirical study. *International Journal of Epidemiology* **31**(1):115-123.

Jüni, P., Witschi, A., Bloch, R., Egger, M. (1999) The hazards of scoring the quality of clinical trials for meta-analysis. *Journal of the American Medical Association*, **282**:1054-1060.

Kalauokalani, D., Cherkin, D.C., Sherman, K.J., Koepsell, T.D., Deyo, R. A. (2001) Lessons from a back pain trial. *Spine*, **26**(13):1418-1424.

Katz, J. N., Barrett, J., Mahomed, N. N., Baron, J. A., Wright, R. J., Losina, E. (2004) Association between hospital and surgeon procedure volume and the outcomes of total knee replacement. *The Journal of Bone and Joint Surgery*, **86-A** (9):1909-1916.

Katz, J. N., Bierbaum, B. E., Losina, E. (2008) Case mix and outcomes of total knee replacement in orthopaedic specialty hospitals. *Medical Care*, **46**:476-480.

Karst, G. M., Boonyawiroj, E. B., Hald, R. D., Harbourne, R. T., Sandstrom, R. W. (1995) Physical therapy intervention and functional ambulation outcomes for patients undergoing total knee arthroplasty. *Issues on aging*, **18** (1):5-9.

Khan, F., Gonzales, S., Hale, T., Turner-Stokes, L. (2008) Multidisciplinary rehabilitation programmes following joint replacement at the hip and knee in chronic arthropathy. *Cochrane database of Systematic Reviews*, Issue 2. Art. No.:CD004957. DOI:10.1002/14651858.CD004957.pub3.

Kim, J. M., Moon, M. S. (1995) Squatting following total knee arthroplasty. *Clinical Orthopaedics and Related Research*, **313**:177-186.

Kim, S., Losina, E., Solomon, D.H., Wright, J., Katz, J.N. (2003) Effectiveness of clinical pathways for total knee and total hip arthroplasty: literature review. *Journal of Arthroplasty*, **18**(1):69-74.

Kjaergard, L. L., Villumson, J., Gluud, C. (2001) Reported methodologic quality and discrepancies between large and small randomised trials in meta-analyses. *Annals of Internal Medicine*, **135**:982-989.

Klaber-Moffett, J., Torgerson, D., Bell-Syer, S., Jackson, D., Llewlyn-Phillips, H., Farrin, Barber, J. (1999). Randomised controlled trial of exercise for low back pain: clinical outcomes, costs, and preferences. *British Medical Journal*, **319**:279-283.

Knipschild, P. (1994) Systematic reviews: some examples. *BMJ*, **309**:719-721.

Kolarz, G., Scherak, O., El-Shohoumi, M., Wottawa, A. (1999) Ergebnisse stationärer rehabilitationsmassnahmen nach knieendoprothesenimplantation - ergebnisse nach einem jahr (Results of inpatient rehabilitation measures after implantation of knee endoprotheses – Results after one year). *Aktuelle Rheumatologie*, **24** (1):22-26.

Koes, B. W., Hoving, J. L. (1998) The value of the randomized clinical trial in the field of physiotherapy. *Manual Therapy*, **3**(4):179-186.

KOOS User's Guide (2003) <http://www.koos.nu/> [accessed 13/05/09].

Kramer, J. F., Speechley, M., Bourne, R., Rorabeck, C., Vaz, M. (2003) Comparison of clinic- and home –based rehabilitation programs after total knee arthroplasty. *Clinical Orthopaedics and Related Research*, **410**:225-234.

Kuettner, K., Goldberg, V.M. (1995) Introduction. In Kuettner, K., Goldberg, V.M., editors. *Osteoarthritic disorders*. Rosemount (IL): American Academy of Orthopaedic Surgeons, p.xxi-v.

Cited by

Brandt, K., Dieppe, P., Radin, E. L. (2008) Etiopathogenesis of osteoarthritis. *Rheumatic Disease Clinics of North America*, **34**:531-559.

Kumar, P. J., McPherson, E. J., Dorr, L. D., Wan, Z., Baldwin, K. (1996) Rehabilitation after total knee arthroplasty a comparison of 2 rehabilitation techniques. *Clinical Orthopaedics and Related Research*, **331**:93-101.

Kunz, R., Oxman, D. (1998) The unpredictability paradox: review of empirical comparisons of randomised and non-randomised clinical trials. *British Medical Journal*, **317**:1185-1190.

Kuster, M. S. (2002) Exercise recommendations after total joint replacement a review of the current literature and proposal of scientifically based guidelines. *Sports Medicine*, **32** (7):433-445.

Landis, J. R., Koch, G. G. (1997) The measurement of observer agreement for categorical data. *Biometrics*, **33**:159-174.

Lang, C. E. (1998) Comparison of 6- and 7-day physical therapy coverage on length of stay and discharge outcome for individuals with total hip and knee arthroplasty. *Journal of Orthopaedic and Sports Physical Therapy*, **28**(1):15-22.

Lam, L. O., Swift, S., Shakespeare, D. (2003) Fixed flexion deformity after knee arthroplasty. What happens in the first 12 months after surgery and can a poor outcome be predicted? *The Knee*, **10**:181-185.

- Lamb, S. E., Frost, H. (2003) Recovery of mobility after knee arthroplasty. *The Journal of Arthroplasty*, **18** (5):575-582.
- Lamb, S. E., Morse, R. E., Grimley Evans, J. (1995) Mobility after proximal femoral fracture: the relevance of leg extensor power, postural sway and other factors. *Age and Aging*, **24**:308-314.
- Lenssen, A.F., Crijns, Y.H.F., Waltjé, E.M.H., van Steyn, M.J.A., Geesink, R.J.T., van den Brandt, P.A., de Bie, R.A. (2006) Efficiency of immediate postoperative inpatient physical therapy following total knee arthroplasty: an rct. *BMC Musculoskeletal Disorders*, **7**:71.
- Lewis, S., Clarke, M. (2001) Forest plots: trying to see the wood and the trees. *British Medical Journal*, **322**:1479-1480.
- Licciardone, J. C., Stoll, S. T., Cardarelli, K. M., Gamber, R. G., Swift, Jr, J. N., Winn, W. B. (2004) A randomized controlled trial of osteopathic manipulative treatment following knee or hip arthroplasty. *Journal of the American Osteopathic Association*, **104** (5):193-202.
- Lieberman, J. R., Dorey, F., Shekelle, P., Schumacher, L., Thomas, B. J., Kilgus, D. J., Finerman, G. A. (1996) Differences between patients' and physicians' expectations of outcome after total hip arthroplasty. *The Journal of Bone and Joint Surgery*, **78-A** (6):835-838.
- Lilford, R. J., Jackson, J. (1995) Equipoise and the ethics of randomisation. *Journal of the Royal Society of Medicine*, **88**:552-559.
- Lilford, R. J., Thornton, J. G., Braunholtz, D. (1995) Clinical trials and rare diseases: a way out of a conundrum. *British Medical Journal*, **311**:1621-1625.
- Lin, Y.C., Davey, R. C., Cochrane, T. (2001) Tests for physical function of elderly with knee & hip osteoarthritis. *Scand J Med Sci Sports*, **11**:280-6.
- Lingard, E.A., Berven, S., Katz, J.N., Kinemax outcomes group (2000) Management and care of patients undergoing total knee arthroplasty: variations across different health care settings. *Arthritis Care Research*, **13**(3):129-136.
- Liow, R., Y. L., Walker, K., Wajid, M. A., Bedi, G., Lennox, C. M. E., (2003) Functional rating for knee arthroplasty: comparison of three scoring systems. *Orthopaedics*, **26**:143-149.
- Llewellyn-Thomas, H. A., McGreal, M. J., Thiel, E. C., Fine, S., Erlichman, C. (1991) Patients' willingness to enter clinical trials: measuring the association with perceived benefit and preference for decision participation. *Social Science and Medicine*, **32**(1):35-42.
- Lord, S. R., Murray, S. M., Chapman, K., Munro, B., Tiedemann, A. (2002) Sit-to-stand performance depends on sensation, speed, balance, and psychological status in

addition to strength in older people. *Journals of Gerontology A Biol Sci Med Sci*, **57A**:M539-43.

McGinn, T., Wyer, P. C., Newman, T. B., Keitz, S., Leipzig, R., Guyatt, G. for the Evidence-Based Medicine Teaching Tips Working Group (2004) Tips for learners of evidence-based medicine:3. Measures of observer variability (kappa statistic). *Canadian Medical Association Journal*, **171** (11):1369-1373.

McLean, S. A. M. (1995) Research ethics committees: Principles and proposals. *Health Bulletin*, **53**(5):243-248.

McPherson, K. (1994) The best and the enemy of the good: randomised controlled trials, uncertainty, and assessing the role of patient choice in medical decision making. *Journal of Epidemiology and Community Health*, **48**:6-15.

Main, E. (2003) Advantages and limitations of systematic reviews in physical therapy. *Cardiopulmonary Physical Therapy Journal*, **14** (2): 24-27.

Mancuso, C.A., Ranawat, C.S., Esdaile, J.M., Johanson, N.A., Charlson, M.E. (1996) Indications for total hip and total knee arthroplasties: results of orthopaedic surveys. *Journal of Arthroplasty*, **11**(1):34-46.

Mansell, J. P., Collins, C., Bailey, A. J. (2007) Bone, not cartilage, should be the major focus in osteoarthritis. *Nature Clinical Practice Rheumatology*, **3**(6):306-307.

Mant, D. (1999) Can randomised trials inform clinical decisions about individual patients? *The Lancet*, **353**:743-746.

Messier, S. P. (2008) Obesity and osteoarthritis: disease genesis and nonpharmacologic weight management. *Rheumatic Disease Clinics of North America*, **34**:713-729.

Milne, S., Brosseau, L., Robinson, V., Noel, M. J., Davis, J., Drouin, H., Wells, G., Tugwell, P. (2003) Continuous passive motion following total knee arthroplasty. *The Cochrane Database of Systematic Reviews*, Issue 2. Art. No.: CD004260. DOI: 10.1002/14651858. CD004260.

Minns Lowe, C.J., Barker, K. L., Dewey, M., Sackley, C. M. (2007) Effectiveness of physiotherapy exercise after knee arthroplasty for osteoarthritis: systematic review and meta-analysis of randomised controlled trials. *British Medical Journal*; **335**:812-815.

Minns Lowe, C.J., Barker, K. L., Dewey, M., Sackley, C. M. (2009) Effectiveness of physiotherapy exercise following hip arthroplasty for osteoarthritis: a systematic review of clinical trials. *BMC Musculoskeletal Disorders*, **10**:98.

Mitchell, C., Walker, J., Walters, S., Morgan, A.B., Binns, T., Mathers, N. (2005) Costs and effectiveness of pre- and post-operative home physiotherapy for total knee replacement. *Journal of Evaluation in Clinical Practice*, **11**(3):283-292.

- Mizner, R.L., Petterson, S.C., Snyder-Mackler, L. (2005) Quadriceps strength and the time course of functional recovery after total knee arthroplasty. *Journal of Orthopaedic and Sports Physical Therapy*, **35**:424-436.
- Mockford, B. J., Beverland, D. E. (2004) Does an outpatient physiotherapy regime improve range of knee motion after total knee arthroplasty; a prospective study. *JBJS*, **86-B**: Suppl III:128.
- Mockford, B. J., Thompason, N. W., Humphreys, P., Beverland, D. E. (2008) Does a standard outpatient physiotherapy regime improve the range of motion after primary total knee arthroplasty. *The Journal of Arthroplasty*, **23** (8):1110-1114.
- Moffet, H., Collet, J-P., Shapiro, S. H., Paradis, G., Marquis, F., Roy, L. (2004) Effectiveness of intensive rehabilitation on functional ability and quality of life after first total knee arthroplasty: a single blind randomized controlled trial. *Archives of Physical Medicine and Rehabilitation*, **85**:546-556.
- Moher, D (2008) Many reviews are systematic but some are more transparent and completely reported than others. *PLoS Medicine*, **4**(3):e147.
- Moher, D., Cook, D. J., Eastwood, S., Olkin, I., Rennie, D., Stroup, D. F., for the Quorum group (1999) Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. *The Lancet*, **354**:1896-1900.
- Moher, D., Fortin, P., Jadad, A. R., Jüni, P., Klassen, T., Le Lorier, J., Liberati, A., Linde, K., Penna, A. (1996) Completeness of reporting of trials published in languages other than English: implication for conduct and reporting of systematic reviews. *Lancet* **347**:363-366.
- Moher, D., Pham, B., Jones, A., Cook, D.J., Jada, A. R., Moher, M., Tugwell, P., Klassen, T. (1998) Does quality of reports of randomised trials affect estimates of intervention quality efficacy reported in meta-analyses? *The Lancet*, **352**:609-613.
- Moher, D., Schulz, K. F., Altman, D. G. for the CONSORT group (2001) The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *Annals of Internal Medicine*, **134**(8):657-662.
- Montgomery, F., Eliasson, M. (1996) Continuous passive motion compared to active physical therapy after knee arthroplasty – similar hospitalization times in a randomised study of 68 patients. *Acta Orthopaedica Scandinavica*, **67**(1):7-9.
- Montori, V.M., Bhandari, M., Devereaux, P.J., Manns, B.J., Ghali, W.A., Guyatt, G.H. (2002) In the dark: the reporting of blinding status in randomized controlled trials. *Journal of Clinical Epidemiology*, **55**(8):787-790.
- Montori, V. M., Swiontkowski, M. F., Cook, D. J. (2003) Methodologic Issues in systematic reviews and meta-analyses. *Clinical Orthopaedics and Related Research*, **413**: 43-54.

Møller, A. M., Villebro, N., Pedersen, T., Tønnesen, H. (2002) A preoperative smoking intervention decreased postoperative complications in elective knee or hip replacement. *Lancet*, **359**:114-117.

Moore, A., Jull, G. (2006) The systematic review of systematic reviews has arrived! *Manual Therapy*, **11**:91-92.

MRC (2000) A framework for development and evaluation of rcts for complex interventions to improve health. [http://www.mrc.ac.uk/pdf-mrc\\_cpr.pdf](http://www.mrc.ac.uk/pdf-mrc_cpr.pdf) [Accessed 12/09/2005].

Mullis, M., Dzeidzic, K.S., Lewis, M., Minns Lowe, C. J., Main, C. J., Watson, P.J., Hay E. M. (2006) Validation of Complex Interventions in a Low Back Pain Trial: Selective video analysis cross-referenced to clinical case notes. *Contemporary Clinical Trials*, **27**(5): 404-412.

Mulrow, C.D. (1994) Rational for systematic reviews. *British Medical Journal*, **309**:597.

Munin, M. C., Rudy, T. E., Glynn, N. W., Crossett, L. S., Rubash, H. E. (1998) Early inpatient rehabilitation after elective hip and knee arthroplasty. *Journal of the American Medical Association*, **279**(11):847-852.

Murray, D. W., Britton, A. R., Bulstrode, C. J. K. (1997) Loss to follow-up matters. *The Journal of Bone and Joint Surgery*, **79-B** (2):254-257.

Murray, D. W., Fitzpatrick, R., Rogers, K., Pandit, H., Beard, D. J., Carr, A. J., Dawson, J. (2007) The use of the Oxford hip and knee scores. *Journal of Bone and Joint Surgery*, **89-B**:1010-1014.

National Institute for Health and Clinical Excellence (2008). *Osteoarthritis The care and management of osteoarthritis in adults NICE clinical guideline 59*. National Institute for Health and Clinical Excellence: London.

National Joint Registry for England and Wales 4th Annual Clinical Report: [http://www.njrcentre.org.uk/documents/reports/annual/4th/4736\\_NJR\\_Annual%20Report\\_WEB.pdf](http://www.njrcentre.org.uk/documents/reports/annual/4th/4736_NJR_Annual%20Report_WEB.pdf) [accessed 27/2/08].

National Joint Registry for England and Wales 5<sup>th</sup> Annual Report. <http://www.njrcentre.org.uk/NjrCentre/LinkClick.aspx?fileticket=Da4%2b2sUvA%2fI%3d&tabid=86&mid=523> [accessed 6<sup>th</sup> May 2009].

Naylor, J., Harmer, A., Fransen, M., Crosbie, J., Innes, L. (2006) Status of physiotherapy rehabilitation after total knee replacement in Australia. *Physiotherapy Research International*, **11**:35-47.

Naylor, J. M., Harmer, A. R., Heard, R. C. (2008) Severe other joint disease and obesity independently influence recovery after joint recovery after joint replacement surgery: an observational study. *Australian Journal of Physiotherapy*, **54**:57-64.

- Noble, P. C., Gordon, M. J., Weiss, J. M., Reddix, R. N., Conditt, M. A., Mathis, K. B. (2005) Does total knee replacement restore normal knee function? *Clinical Orthopaedics and Related Research*, **431**:157-165.
- Norkin, C. C., White, D. J. (2003) *Measurement of joint motion a guide to goniometry* (third edition). F.A.Davis Company: Philadelphia.
- Noseworthy, J.H., Ebers, G.C., Vandervoort, M.K., Farquhar, R.E., Yetisir, E., Roberts, R. (1994) The impact of blinding on the results of a randomised placebo controlled multiple sclerosis clinical trial. *Neurology*, **44**:16-20.
- Nyberg, B., Kreuter, M. (2002) Sjukgymnastisk gruppträning jämfört med hemträning för patienter opererade med total höftledsplastik (Physiotherapist-led group training compared to home training for patients receiving total hip replacements). *Nordisk Fysioterapi*, **6**:82-88.
- Oakley, A. (1998) Experimentation and social interventions: a forgotten but important history. *British Medical Journal*, **317**:1239-1242.
- Oldmeadow, L. B., McBurney, H., Robertson, V. J. (2002) Hospital stay and discharge outcomes after knee arthroplasty: implications for physiotherapy practice. *Australian Journal of Physiotherapy*, **48**:117-121.
- Oppenheim, A. N. (1992) *Questionnaire design, interviewing and attitude measurement*. Second Edition. Pinter Publishers Limited: London.
- Pallant, J. (2007) *SPSS survival manual third edition*. Open University Press: Maidenhead.
- Parker, C., Dewey, M. (2000) Assessing research outcomes by postal questionnaire with telephone follow up. TOTAL study group. Trial of occupational therapy and leisure. *International Journal of Epidemiology*, **29**(6):1065-1069.
- Parvazi, J., Mui, A., Purtill, J. J., Sharkey, P. F., Hozack, W. J., Rothman, R. H. (2007) Total joint arthroplasty: when do fatal or near-fatal complications occur. *The Journal of Bone and Joint Surgery*, **89-A** (1):27-32.
- Peat, G., McCarnery, R., Croft, P. (2001) Knee pain and osteoarthritis in older adults: a review of community burden and current use of health care. *Annals of the Rheumatic Diseases*, **60**:91-97.
- Peersman, G., Laskin, R., Davis, J., Petersen, M. (2001) Infection in total knee replacement. *Clinical Orthopaedics and Related Research*, **392**:15-23.
- Peto, V., Coulter, A., Bond, A. (1993) Factors affecting general practitioners' recruitment of patients into a prospective study. *Family Practice*, **10**:207-211.
- Peto, R., Baigent, C. (1998) Trials: the next 50 years Large scale randomised evidence of moderate benefits. *British Medical Journal*, **317**:1170-1171.



Petterson, S.C., Mizner, R.L., Stevens, J.E., Rasis, L., Bodenstab, A., Newcomb, W., Snyder-Mackler, L. (2009) Improved function from progressive strengthening interventions after total knee arthroplasty: a randomised clinical trial with an imbedded prospective cohort. *Arthritis & Rheumatism*, **61**(2):174-183.

Pettigrew, M. (2003) Why certain systematic reviews reach uncertain conclusions. *British Medical Journal*, **326**: 756-758.

Pocock, S. J. (1983) *Clinical trials a practical approach*. John Wiley & Sons: Chichester, UK.

Pollard, B., Dixon, D., Dieppe, P., Johnstone, M. (2009) Measuring the ICF components of impairment, activity limitation and participation restriction: an item analysis using classical test theory and item response theory. *Health and Quality of Life Outcomes*, **7**:41.

Pollard, B., Johnston, M., Dieppe, P. (2006) What do osteoarthritis health outcome instruments measure? Impairment, activity limitation, or participation restriction? *Journal of Rheumatology*, **33**(4):757-763.

Poolman R.W., Struijs, P.A.A., Krips, R., Sierevelt, I.N., Marti, R.K., Farrokhyar, F., Bhandari, M. (2007) Reporting of outcomes in orthopaedic randomized trials: does blinding of outcome assessors matter? *The Journal of Bone and Joint Surgery*, **98-A** (3):550-558.

Pope, C., Ziebland, S., Mays, N. (2000) Qualitative research in health care analysing qualitative data. *British Medical Journal*, **320**:114-116.

Prescott, R.J., Counsell, C.E., Gillespie, W.J., Grant, A.M., Russell, I.T., Kiauka, S., Colthart, I.R., Ross, S., Shepherd, S.M., Russell, D. (1999) Factors that limit the quality, number and progress of randomised controlled trials. *Health Technology Assessment*, **3** (20) p. 74.

Pringle, M., Churchill, R. (1995) Randomised controlled trials in general practice. *British Medical Journal*, **311**: 1382-1383.

R Development Core Team (2006). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing: Vienna.

Rajan, R. A., Pack, Y., Jackson, H., Gillies, C., Asirvatham, R. (2004) No need for outpatient physiotherapy following total knee arthroplasty a randomised trial of 120 patients. *Acta Orthop Scand*, **75** (1):71-73.

Rand, J. A., Trousdale, R. T., Ilstrup, D. M., Harmsen, W. S. (2003) Factors affecting the durability of primary total knee prostheses. *The Journal of Bone and Joint Surgery*, **85-A**:259-265.

Rees, J.R., Wade, T.J., Levy, D.A., Colford Jr., J.M., Hilton, J.F. (2005) Changes in beliefs identify unblinding in randomized controlled trials: a method to meet CONSORT guidelines. *Contemporary Clinical Trials*, **26**(1):25-37.

- Richmond, J. C. (2008) Surgery for osteoarthritis of the knee. *Rheumatic Disease Clinics of North America*, **34**:815-825.
- Ritter, M. A., Gandolf, V.S., Holston, K. S. (1989) Continuous passive motion versus physical therapy in total knee arthroplasty. *Clinical Orthopaedics and Related Research*, **244**:239-243.
- Rivero-Arias, O., Campbell, H., Gray, A., Fairbank, J., Frost, H., Wilson-MacDonald, J. (2005) Surgical stabilisation of the spine compared with a programme of intensive rehabilitation for the management of patients with chronic low back pain: cost utility analysis based on a randomised controlled trial. *British Medical Journal*, **330**:1239-1243.
- Robertson, S., Frost, H., Doll, H., O'Connor, J. J. (1998) Leg extensor power and quadriceps strength: an assessment of repeatability in patients with osteoarthritic knee. *Clinical Rehabilitation*, **12** (2):120-126.
- Robertsson, O., Dunbar, M. J. (2001) Patient satisfaction compared with general health and disease-specific questionnaires in knee arthroplasty patients. *The Journal of Arthroplasty*, **16** (4):476- 482.
- Rogers, J., Shepstone, L., Dieppe, P. (2004) Is osteoarthritis a systemic disease of bone? *Arthritis and Rheumatism*, **50**(2):452-457.
- Roland, M., Torgerson, D. J. (1998) Understanding controlled trials: What are pragmatic trials. *British Medical Journal*, **316**:285.
- Roos, E. M. (2003) Effectiveness and practice variation of rehabilitation after joint replacement. *Current Opinion in Rheumatology*, **15**:160-162.
- Roos, E. M., Lohmander, L. S. (2003) The knee injury and osteoarthritis outcome score (KOOS): from joint injury to osteoarthritis. *Health and Quality of Life Outcomes*, **1**(1):64.
- Roos, E.M., Toksvig-Larsen, S (2003) Knee injury and osteoarthritis outcome score (KOOS) – validation and comparison to the WOMAC in total knee replacement. *Health and Quality of Life Outcomes*, **1**:17.
- Rothstein, J. M., Miller, P. J., Roettger, R. F. (1983) Goniometric reliability in a clinical setting elbow and knee measurements. *Physical Therapy*, **63**:1611-1615.
- Rothwell, P. M. (2005) External validity of randomised controlled trials: “to whom do the results of this trial apply?”. *The Lancet*, **365**:82-93.
- Rowe, P. J., Myles, C. M., Walker, C., Nutton, R. (2000) Knee joint kinematics and other functional activities measured using flexible electrogoniometry: how much knee motion is sufficient for normal daily life? *Gait and Posture*, **12**:143-155.

Sackett, D. L. (2004) Turning a blind eye why we don't test for blindness at the end of our trials. *British Medical Journal*, **328**:1136.

Sackett, D. L. (2007) Commentary: Measuring the success of blinding in RCTs: don't, must, can't or needn't? *International Journal of Epidemiology*, **36**:664-665.

Sacks, H. S., Berrier, J., Reitman, D., Ancona-Berk, V. A., Chalmers, T. C. (1987) Meta-analyses of randomised controlled trials. *The New England Journal of Medicine*, **316**:450-455.

Santaguida, P.L., Hawker, G. A., Hudak, P. L., Glazier, R., Mahomed, N. N., Kreder, H. J., Coyte, P. C., Wright, J. G. (2008) Patient characteristics affecting the prognosis of total hip and knee joint arthroplasty: a systematic review. *Canadian Journal of Surgery*, **51** (6):428-436.

Sarhmann, S.A. (2002) *Diagnosis and treatment of movement impairment syndromes*. Mosby, Inc: St. Louis.

Schaubert, K. L., Bohanon, R. W. (2005) Reliability and validity of three strength measures obtained from community-dwelling elderly persons. *Journal of Strength and Conditioning Research*, **19** (3):717-720.

Schulz, K.F. (2000) Assessing allocation concealment and blinding in randomised controlled trials: why bother? *EBM Notebook*, **5**:36-37.

Schulz, K.F., Chalmers, I., Altman, D.G. (2002) The landscape and the lexicon of blinding in randomized trials. *Annals of Internal Medicine* **136** (3):254-259.

Schulz, K.F., Grimes, D.A. (2002) Blinding in randomised trials: hiding who got what. *The Lancet*, **359**(9307):696-700.

Scottish Arthroplasty Project Annual Report 2008

[http://www.arthro.scot.nhs.uk/Reports/Scottish\\_Arthroplasty\\_Project\\_Report\\_2008.pdf](http://www.arthro.scot.nhs.uk/Reports/Scottish_Arthroplasty_Project_Report_2008.pdf) [accessed 18/05/2009].

Shamley, D. R., Barker, K., Simonite, V., Beardshaw, A. (2005) Delayed versus immediate exercises following surgery for breast cancer: a systematic review. *Breast Cancer Research and Treatment*, **90**:263-271.

Shields, R. K., Enloe, L. J., Evans, R. E., Smith, K. B., Steckel, S. D. (1995) Reliability, validity, and responsiveness of functional tests in patients with total joint replacement. *Physical Therapy*, **75**, 3:169-79.

Showalter, A., Burger, S., Salyer, J. (2000) Patients' and their spouses' needs after total joint arthroplasty: a pilot study. *Orthopaedic Nursing*, **19**(1):49-58.

Shrier, I. (2003) Cochrane reviews: new kids on the block. *Br J Sports Med*, **37**:473-474.

Shrier, I. (2005) Power, reliability and heterogeneous results. *PLoS Medicine*, **2** (11); e386.

Sibbald, B., Roland, M. (1998) Understanding controlled trials: why are randomised controlled trials important? *British Medical Journal*, **316**:201.

Silva, M., Shepherd, E. F., Jackson, W. O., Pratt, J. A., McClung, C. D., Schmalzried, T. P. (2003) Knee strength after total knee arthroplasty, *The Journal of Arthroplasty*, **18** (5):605-611.

Sim, J., Reid, N. (1999) Statistical inference by confidence intervals: issues of interpretation and utilization. *Physical Therapy*, **79**(2):186-195.

Sim, J., Wright, C. (2000) *Research in health care concepts, designs and methods*. Stanley Thornes (Publishers) Ltd: Cheltenham.

Smidt, N., de Vet, H.C.W., Bouter, L.M., Dekker, J. (2005) Effectiveness of exercise therapy: a best-evidence summary of systematic reviews. *Australian Journal of Physiotherapy*, **51**:71-85.

Solomon, D.H., Chibnik, L. B., Losina, E., Huang, J., Fossel, A. H., Husni, E., Katz, J. N. (2006) Development of a preliminary index that predicts adverse events after total knee replacement. *Arthritis & Rheumatism*, **54**(5):1536-1542.

Soohoo, N. F., Lieberman, J. R., Ko, C. Y., Zingmond, D. S. (2006) Factors predicting complications rates following total knee replacement. *The Journal of Bone and Joint Surgery*, **88-A** (3):480-485.

Stanley D. Pilot randomised controlled trial of an early rehabilitation class versus outpatient physiotherapy following discharge from hospital after primary knee arthroplasty. National Research Register Document – N0109133246. [Accessed 24/4/07].

Steadman, J., Donaldson, N., Kalra, L. (2003) A randomized controlled trial of an enhanced balance training program to improve mobility and reduce falls in elderly patients. *Journal of the American Geriatrics Society*, **51**(6):847-852.

Stevens, J. E. (2002) Physiologically-based rehabilitation for patients after total knee arthroplasty. University of Delaware, PhD thesis.

Stewart, R. DHSSPSNI, (6 January 2009). Re: total knee replacement statistics. E-mail personal communication to C. Minns Lowe.

Su, F. C., Lai, K. A., Hong, W. H. (1998) Rising from chair after total knee arthroplasty. *Clinical Biomechanics*, **13**:176-181.

Sultan, P.G., Schule, S., Li, G., Rubash, H.E. (2003) Optimizing flexion after total knee arthroplasty: advances in prosthetic design. *Clinical Orthopaedics and Related Research*; **416**:167-73.

Sun, Y., Stürmer, T., Günther, K. P., Brenner, H. (1997) Reliability and validity of clinical outcome measurements of osteoarthritis of the hip and knee – a review of the literature. *Clinical rheumatology*, **16** (2):185-198.

Swales, J. (1998) Research and development in the NHS. *Journal of the Royal Society of Medicine*, **91** Suppl 36:18-20.

Thompson, S. G. (1994) Systematic review: why sources of heterogeneity in meta-analysis should be investigated. *BMJ*, **309**:1351-1355.

Trudelle-Jackson, E. (2002) Balance and postoperative lower extremity joint replacement. *Orthopaedic Physical Therapy Clinics of North America*, **11** (1):101-110.

Trudelle-Jackson, E., Smith, S. (2004) Effects of a late-phase exercise program after total hip arthroplasty: a randomised controlled trial. *Archives of Physical Medicine and Rehabilitation*, **85**:1056-1062.

Tum-Sugden, A. (1976) Krankengymnastische behandlung in verbindung mit eis bei prothetischem kniegelenkersatz (Remedial exercises with application of ice in cases of prosthetic replacement of the knee joint). *Krankengymnastik*, **28**(3):91-92.

Turner, J.A., Jensen, M.P., Warms, C.A., Cardenas, D.D. (2002) Blinding effectiveness and association of pre-treatment expectations with pain improvement in a double blinded randomised controlled trial. *Pain*, **99**(1):91-99.

Uebersax, J. (2002) kappa coefficients  
<http://ourworld.compuserve.com/homepages/jsuebersax/kappa.htm>  
[Accessed 2<sup>nd</sup> November 2006].

Ulreich, A., Kullich, W., Klein, G., Ramach, W. (1997) Ergebnisse eines multidisziplinären rehabilitationskonzepts nach kniegelenkersatz (Results of a multidisciplinary rehabilitation concept after total knee replacement) *Aktuelle Rheumatologie*, **22**(6):211-216.

Ulrich, S.D., Bhave, A., Marker, D.R., Seyler, T.M., Mont, M.A. (2007) Focused rehabilitation treatment of poorly functioning total knee arthroplasties. *Clinical Orthopaedics and Related Research*, **464**:138-145.

van der Windt, D. (1997) Practical aspects of conducting a pragmatic randomised trial in primary care. Chapter 8 in *Shoulder Disorders in Primary Care*, p123-134. Institute for Research in Extramural Medicine (EMGO Institute) Vrije Universiteit: The Netherlands, Phd Thesis.

van Peppen RPS, Kwakkel G, Wood-Dauphinee S, Hendriks HJM, Van der Wees PhJ, Dekker J. (2004) The impact of physical therapy on functional outcomes after stroke: what's the evidence? *Clinical Rehabilitation*, **18**:833-862.

Vickers, A. J., Altman, D. G. (2001) Analysing controlled trials with baseline and follow up measurements. *British Medical Journal*, **323**:1123-1124.

- Viera, A. J., Bangdiwala, S. I. (2007). Eliminating bias in randomised controlled trials: importance of allocation concealment and masking. *Family Medicine*, **39**(2):132-137.
- Wade, D. (2005) Describing rehabilitation interventions. *Clinical Rehabilitation*, **19**: 811-818.
- Walsh, M., Woodhouse, L. J., Thomas, S. G., Finch, E. (1998) Physical impairments and functional limitations: a comparison of individuals 1 year after total knee arthroplasty with control subjects. *Physical Therapy*, **78**(3):248-258.
- Walter, S.D., Awasthi, S., Jeyaseelan, L. (2005) Pre-trial evaluation of the potential for unblinding in drug trials: A prototype example. *Contemporary Clinical Trials*, **26**(4): 459-468.
- Ward, E., King, M., Lloyd, M., Bower, P., Friedli, K. (1999) Conducting randomized trials in general practice: methodological and practical issues. *British Journal of General Practice*, **49**:919-922.
- Waters, E. A. (1974) Physical therapy management of patients with total knee replacement. *Physical Therapy*, **54**(9):936-942.
- Watkins, M. A., Riddle, D. L., Lamb, R. L., Personius, W. J. (1991) Reliability of goniometric measurements and visual estimates of knee range of motion obtained in a clinical setting. *Physical Therapy*, **71**:90-96.
- Westby, M.D., Kennedy, D., Jones, D.L., Jones, A., Doyle-Waters, M., Backman, C. (2008) Post-acute physiotherapy for primary total knee arthroplasty. (Protocol) *Cochrane Database of Systematic Reviews*, Issue 2. Art. No.:CD007099. DOI:10.1002/14651858. CD007099.
- Whitehead, C., Miller, M., Crotty, M. (2003) Falls in community-dwelling older persons following hip fracture: impact on self-efficacy, balance and handicap. *Clinical Rehabilitation*, **17**:899-906.
- Whitehouse, S. L., Blom, A. W., Taylor, A. H., Pattison, G. T. R., Bannister, G. C. (2005) The Oxford knee score; problems and pitfalls. *The Knee*, **12**:287-291.
- Wolinsky, F. D., Miller, D., K., Andresen, E. M., Malmstrom, T. K., Miller, S. P. (2005) Reproducibility of physical performance and physiologic assessments. *Journal of Aging and Health*, **17** (2):111-124.
- Woolhead, G. M., Donovan, J. L., Dieppe, P. A. (2005) Outcomes of total knee replacement: a qualitative study. *Rheumatology*, **44**:1032-1037.
- Worland, R. L., Arredondo, J., Angles, F., Lopez-Jimenez, F., Jessup, D. E. (1998) Home continuous passive motion machine versus professional physical therapy following total knee replacement. *The Journal of Arthroplasty*, **13**(7):784-787.

World Health Organisation (2002) *Towards a common language for Functioning, Disability and Health ICF*. World Health Organisation: Geneva.

Wright, J.G., Rudical, S., Feinstein, A.R. (1994) Ask patient what they want. Evaluation of individual complains before total hip arthroplasty. *Journal of Bone and Joint Surgery*, **76-B**:229-234.

Wylde , V.,Blom, A., Dieppe, P., Hewlett, S., Learmonth, I. (2008) Return to sport after joint replacement. *The Journal of Bone and Joint Surgery*, **90-B**:920-923.

Wylde, V., Dieppe, P., Hewlett, S., Learmonth, I. D. (2007) Total knee replacement: is it really an effective procedure for all? *The Knee*, **14**:417-423.

Xie, F., Thumboo, J., Fong, K-Y., Lo, N-N., Yeo, S-J., Li, S-C. (2006) Are they relevant? A critical evaluation of the international classification of functioning, disability, and health core sets for osteoarthritis from the perspective of patients with knee osteoarthritis in Singapore. *Annals of Rheumatic Diseases*, **65**:1067-1073.

Yaffee, R. A. (2003) Common correlation and reliability analysis with SPSS for windows. <http://www.nyu.edu/its/statistics/Docs/correlate.html> [Accessed 25/10/2006].

Zwarenstein, M., Treweek, S., Gagnier, J.J., Altman, D. G., Tunis, S., Haynes, B., Oxman, A. D., Moher, D and the CONSORT and Pragmatic Trials in Healthcare (Practihc) groups (2008). Improving the reporting of pragmatic trials: an extension of the CONSORT statement. *British Medical Journal*, **337**:a2390,1-8.

**APPENDIX I. Permission to reproduce Figure 1.**



## **APPENDIX II. LANGUAGE BIAS: ITS IMPLICATIONS FOR A PROPOSED SYSTEMATIC REVIEW.**

This assignment was written by Catherine Minns Lowe and submitted as part of the Systematic Reviews Module, Department of Continuing Education, University of Oxford 01/07/2005. It was awarded Grade A.

### **Introduction.**

This assignment addresses the question: should our planned review (Sub Appendix I), evaluating the effectiveness of post-operative physiotherapy exercise following hospital discharge for patients undergoing elective total hip/knee arthroplasty, be restricted to the inclusion of English language trial reports, or should all appropriate reports be considered regardless of their language?

### **Background.**

The inclusion of an unbiased selection of relevant studies in a systematic review is central to its validity<sup>1</sup>. Many systematic reviews exclude non English Language trial (NELT) reports<sup>1-7</sup>. Excluding NELTs may introduce bias, termed language bias<sup>1</sup>, however, including them may be problematic<sup>1,3</sup>. It was therefore decided to ascertain how language bias might effect our review and whether NELTs should be included.

### **Options.**

Information was obtained (Sub Appendix II) and considered. Four possible options were identified. Firstly, exclude NELTs from our review. Secondly, include all papers identified by the search strategy, regardless of language of publication. Thirdly, screen all abstracts identified by the search strategy and only obtain papers meeting pre-specified criteria. The final option was to ignore the issue. These options will now be explored.

#### **1. Effects of Exclusion.**

Exclusion of NELTs can affect the results of meta-analyses<sup>2</sup>, excluding around 20%<sup>1,3</sup> of trials from Medline searches and more for databases relevant to our search such as Embase<sup>3</sup> and CINAHL. Although exclusion generally has little effect on treatment effect estimates, the size and direction of any differences are unpredictable and precision may be reduced<sup>1,5</sup>. This is relevant for our review since treatment effects sizes in trials are usually small and even small changes in meta-analysis results may have clinical import. Evidence exists that authors are likely to report trials with statistically significant results in English, although this trend may be decreasing<sup>6</sup>. Again this could effect our review since we may include exercises in our intervention believing they are effective when they are actually less / not effective. Additionally the number of papers included in a review is relevant<sup>2</sup>. If a review includes a large number of papers then excluding one or two NELTs may not make any significant difference to the results. However for our review, since physiotherapy research is often limited or inadequate<sup>8</sup>, any exclusions might alter the results in a significant way; as demonstrated in small reviews, which may even reach a different conclusion if NELTs are excluded<sup>1,3</sup>, with consequential influence upon clinical practice. Or, where reviews fail to reach a conclusion when a conclusion may have been possible<sup>7</sup>. For our review, even though huge heterogeneity prevents meta-analysis, NELTs remain important since all available evidence is required to assist the development of

a trial intervention (Appendix I). Furthermore, excluding such papers casts doubt upon the thoroughness of our review<sup>1-2</sup> and our decision making<sup>7</sup>. Greater confidence may be inspired by their inclusion<sup>1</sup> since language inclusiveness indicates a higher quality review<sup>9</sup>.

## **2. Effects of Inclusion.**

The inclusion of NELTs may also adversely effect our review. The methodological quality of NELTs tends to be poorer<sup>1,5</sup>, although these differences in quality may be small<sup>4,6</sup>, and it might be expected that these differences may also be unpredictable. For our review this is a concern. Physiotherapy trials have traditionally been of low quality<sup>8</sup> even in English. Plus, we believe physiotherapy practice varies widely across countries so practice in some countries may be inappropriate to inform practice in England. Including NELTs can also create difficulties when findings have been reported in both English and non English<sup>6,8</sup> since the duplication of findings can effect results. We have identified one duplicate trial in English and German.

If NELTs are reported less fully than those published in English then it also might be appropriate not to include them<sup>4</sup>. We plan to screen trials for quality prior to inclusion. Additionally, work on language bias has been limited to several languages and languages / countries do not seem to respond similarly to language bias<sup>1,4-5</sup>. The NELTs languages involved in our review (Swedish and Russian) have not been researched and so if / how they are likely to be effected by language bias is unknown.

The importance of including / excluding NELTS depends on the review's topic area<sup>6,9</sup>. It has been recommended that for orthopaedic reviews like ours it is important to include NELTs<sup>1,6</sup>, and to search widely since Medline is not able to identify all relevant trials<sup>3,6</sup> which we have done (Appendix I).

Furthermore, including NELTs can add considerably to the costs and time required to undertake the review and delay the review's conclusion<sup>1,3,9</sup>. Although, with the development of the Cochrane Central Register of Controlled Trials (CENTRAL), locating NELTs should get easier<sup>10</sup>, locating physiotherapy trials is currently time-consuming. CENTRAL identified none of our three identified NELTs. Our costs and time are constrained with the trial starting 01/10/2005.

If we include NELTs we need to address who/how should translations be performed? Reviewers have used translations by themselves, colleagues or professional translators<sup>2,4</sup>. Sandford (1996)<sup>11</sup> opines that specialist subject knowledge, plus fluency, is required to obtain an accurate translation which incorporates how medical practice, language, abbreviations and cultural differences vary across countries with only professional medical translators being used. This approach would add significantly to our costs.

## **3. Screening.**

A further approach is to have an abstract or article translated "to the degree necessary"<sup>3</sup> to determine whether it meets the review inclusion criteria and then take further action as appropriate. This approach averts the decision to include / exclude all NELTS and makes an individual decision for each trial report. This would appear to be a pragmatic and cost effective practical approach to relate to our review.

#### **4. Ignore.**

Given the evidence that language bias exists we decided to reject this option. We concur that it “surprising” how often reviews are limited, without discussion, to English language publications<sup>2</sup>.

#### **Decision.**

Although our review is constrained, we believe it is necessary to include all identified trials meeting our criteria, regardless of language. Although meta-analysis will be inappropriate for our review, we feel that the introduction of language restrictions might affect the quality of our future trial intervention; an avoidable risk. We accept Sandford’s (1996)<sup>11</sup> concern regarding the use of non-professional medical translators, however we have budgetary restrictions and need to be pragmatic. We will therefore screen our NELTs with colleagues who come from Russia and Sweden. Any reports appearing pertinent to our review will subsequently be sent for full professional translation. This approach meets our constraints, since the number of NELTS is few, and has been approved by the review team and external experienced allied health professional systematic reviewers.

#### **REFERENCES.**

1. Jüni, P., Holenstein, F., Sterne, J., Bartlett, C., Egger, M. (2002) Direction and impact of language bias in meta-analyses of controlled trials: empirical study. *International Journal of Epidemiology* **31**(1): 115-123.
2. Grégoire, G., Derderian, F., Le Lorier (1995) Selecting the language of the publications included in a meta-analysis: is there a tower of babel bias? *Journal of Clinical Epidemiology* **48**(1): 159-163.
3. Dickersin, K., Scherer, R., Lefebvre, C. (1994) Identifying relevant studies for systematic reviews. *British Medical Journal* **309**: 1286.
4. Moher, D., Fortin, P., Jadad, A. R., Jüni, P., Klassen, T., Le Lorier, J., Liberati, A., Linde, K., Penna, A. (1996) Completeness of reporting of trials published in languages other than English: implication for conduct and reporting of systematic reviews. *Lancet* **347**: 363-366.
5. Egger, M., Jüni, P., Bartlett, C., Holenstein, F., Sterne, J. (2003) How important are comprehensive literature searches and the assessment of trial quality in systematic reviews? Empirical study. *Health Technology Assessment* **7**(1).
6. Eggar, M., Zellweger-Zähner, T., Schneider, M., Junker, C., Lengeler, C., Antes, G. (1997) Language bias in randomised controlled trials published in English and German. *Lancet* **350**: 326-329.
7. Kovacs, F. M., & Abaira, V. (2004) Language bias in a systematic review of chronic pain: how to prevent the omission of non-English publications? *Clinical Journal of Pain* **20**(3): 199-200.

8. Main, E. (2003) Advantages and limitations of systematic reviews in physical therapy, *Cardiopulmonary Physical Therapy Journal* **14**(2): 24-27.
9. Moher, D., Pham, B., Lawson, M. L., Klassen, T. P. (2003) The inclusion of reports of randomised trials published in languages other than English in systematic reviews. *Health Technology Assessment* **7**: No. 41.
10. Clarke, M. (2002) Commentary: searching for trials in systematic reviews: what difference does it make. *International Journal of Epidemiology* **31**(1): 123-124.
11. Sandford, P. (1996) Completeness of reporting of trials published in languages other than English. *Lancet* **347**: 907-908.

### **BIBLIOGRAPHY.**

Cochrane Reviewer's Handbook 4.4.2. Chapter 5.0. Locating and Selecting studies for Reviews. <http://www.cochrane.dk/cochrane/handbook/hbook.htm>

Egger, M., Dickersin, K., Davey Smith, D. (2001) 'Problems and limitations in conducting systematic reviews', in: Egger, M., Davey Smith, G., Altman, D. G., (eds) *Systematic Reviews in Health Care: Meta-analysis in Context*, 2<sup>nd</sup> edition. London: BMJ Publishing Group.

Herrera, A. J. (1999) Language bias discredits the peer-review system. *Nature* **397**: 467.

Sterne, J. A. C., Bartlett, C., Jüni, P., Egger, M. (2000) Do we need comprehensive literature searches? A study of publication and language bias in meta-analyses of controlled trials. *Proceedings of the 3<sup>rd</sup> Symposium on Systematic Reviews: beyond the basics*. Oxford, 3-5 July 2000.

University of Oxford Department for Continuing Education: Continuing Professional Development Centre (2005) *Systematic Reviews* course notes. Course Tutor: Mike Clarke. Oxford: Continuing Professional Development Centre.

### ***SUB APPENDIX I.***

#### **A SUMMARY OF OUR PROPOSED REVIEW.**

**Title:** A systematic review evaluating the effectiveness of post operative physiotherapy exercise following discharge from hospital (in terms of improving function, mobility and strength) for patients undergoing elective total hip or knee arthroplasty (joint replacement).

**Background:** Generally there is a marked and rapidly decreasing length of stay following arthroplasty. Previous work suggests that early discharge after arthroplasty frequently results in persistent muscle weakness and poor functional mobility and that further rehabilitation within primary care may be required to optimise recovery for

these patients. It is not yet known whether such physiotherapy interventions are effective in improving function, mobility, strength or range of movement.

**Purpose:** This review is primarily being undertaken to inform the development of a new complex intervention for use in a randomized controlled trial “Is post discharge physiotherapy effective in improving function and mobility following elective joint arthroplasty in a Diagnostic and Treatment Centre?” The trial is due to start in October 2005. The purpose of the trial is to discover whether patients discharged early following elective joint arthroplasty significantly benefit from additional physiotherapy post discharge, compared to usual care.

**Participants:** The participants for inclusion in this review are patients undergoing elective total hip or knee arthroplasty who have received physiotherapy following discharge from hospital post-operatively.

**Intervention:** The intervention “physiotherapy” is referred to by different names in the literature; namely, physiotherapy, physical therapy, home programme, exercise and rehabilitation. For the purposes of this review the term physiotherapy refers to any exercises or exercise programme provided by physiotherapists/physical therapists during the rehabilitative period after discharge from hospital after joint replacement surgery occurring in the out patient, community or home setting.

**Comparison:** The review will include randomized clinical trials relevant to the review question, regardless of language of publication. Trials will be included if they compare a post discharge physiotherapy exercise intervention compared with either control group, usual care or another form of physiotherapy exercise intervention.

**Outcomes:** The relevant outcomes by which effectiveness is to be measured include the following: Function, Mobility, Muscle strength, Range of motion.

**Location of Studies:** Early searches were conducted prior to this systematic review to begin to identify appropriate search terms. Since these searches indicated that a large number of terms seemed relevant, a broad area search was initially undertaken which was followed by a number of more specific searches to check that as many relevant records as possible were identified. It was accepted that this inclusive approach would identify many papers irrelevant to the review but this approach seemed to offer the most comprehensive method of locating the maximum possible number of trials present. The databases searched were: AMED 1985-; CINAHL 1982-; EMBASE 1974-; KingsFund 1979-; MEDLINE 1966-; The Cochrane library (including Cochrane reviews, Cochrane Central Register of Controlled Trials, DARE); PEDro physiotherapy evidence database; The Department of Health National Research Register

**Terms searched** included:

Hip, knee, replacement, arthroplasty, trial, exercise, physiotherapy and physical therapy, home programme.

**Selection of Studies:** Two reviewers will independently assess the eligibility of identified studies for inclusion into the review against specified criteria. A consensus meeting will be held to discuss discrepancies.

**Data Extraction:** Data regarding the quality of the studies and their results will be independently extracted by two reviewers. A consensus meeting will be held to discuss discrepancies and a third reviewer is available for further discussion.

**Quality:** A form, based upon the CONSORT statement<sup>1</sup> and the CASP guidelines<sup>2</sup> has been developed to extract information for the review. This form includes rating the following aspects: adequate randomisation procedures; adequate concealment; blinding of outcome assessment; attrition; power and size of study; effect size.

**Analysis:** This stage of the protocol is still being developed. A statistician is being consulted regarding the design of this section. It is anticipated that, since the studies included in this view are not likely to be heterogenous, a meta-analysis will be inappropriate and potentially misleading for this review.

1. Altman, D. G., Schulz, K. F., Moher, D., Egger, M., Davidoff, F., Elbourne, D., Gøtzsche, P. C., Lang, T. (2001) The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Annals of Internal Medicine*, **134**: 663-694.
2. Critical Appraisal Skills Programme (accessed 14/05/05) 10 questions to help you make sense of randomised clinical trials. <http://www.phru.nhs.uk/casp/rct.pdf>

### **SUB APPENDIX II.**

To identify information for this assignment on line databases, including Medline, Embase and Cinahl, were searched using the title term “language bias” AND document term “systematic review”. The Cochrane Library Methodology Reviews, Cochrane Reviewer’s Handbook 4.4.2, and the National Coordinating Centre for Health Technology Assessment Monograph Series on CD-ROM were also searched. Relevant papers identified from these searches were obtained and their reference lists and citations were searched for further publications, as were references lists from course notes and the Systematic Reviews Course’s recommended text (see bibliography). This search strategy was not intended to be 100% exhaustive but was designed to provide sufficient relevant and appropriate information for this 1000 word assignment.

## APPENDIX III. 22 Item Checklist. (based on the CONSORT<sup>1</sup> statement and CASP<sup>2</sup> guidelines.)

Authors and Year of Publication:

Item Number	Paper Section and Topic	Descriptor: Does the study include/ mention this?
1	Title & abstract	how participants were allocated to interventions? (eg: “random allocation” “randomized”)
2	Introduction: Background	the scientific background and explanation of rationale included?
3	Methods: Participants	Eligibility criteria
		Method of recruitment
		Setting and locations where data was collected
4	Methods: Interventions	Precise details of the interventions intended for each group and how and when they were actually administered
5	Methods: Objectives	Specified objectives and hypotheses
6	Methods: Outcomes	Clearly defined primary and secondary outcome Measures.
		Please state outcome measures used
		When applicable, any methods used to enhance the quality of measurement (e.g. multiple observation, training of assessors)
7	Methods: Sample size	How sample size was determined and, when applicable, explanation of any interim analyses  PLEASE STATE SAMPLE SIZE
	Methods: randomisation	Does it state that randomisation has occurred?
8	Methods: Randomisation Sequence generation	Method used to generate the random allocation sequence, including details of any restriction (Eg: blocking or stratification)
9	Methods: Randomisation Allocation concealment	Method used to implement the random allocation Sequence, cf: whether the sequence was concealed until interventions were assigned
10	Methods: Randomisation Implementation	Who generated the allocation sequence, who enrolled Patients, and who assigned patients to their groups
11	Methods: Blinding	Were participants blinded to allocation? (Consider that this may not be possible <sup>2</sup> )
		Were those administering the intervention blinded? (Consider that this may not be possible <sup>2</sup> )
		Was outcome assessment blinded?
12	Methods: Statistical methods	Methods used to compare groups for primary outcomes and methods for additional analyses (such as subgroup and adjusted). Please state.

13	Results: Participant flow	Flow of participants through each stage (i.e. number Randomized, receiving intended treatment, completing protocol, analysed. Also, protocol deviations (& reasons) as planned.
14	Results: Recruitment	Dates defining the periods of recruitment and follow-up
15	Results: Baseline data	Baseline demographics and clinical characteristics presented for each group
16	Results: Numbers analysed	No of participants in each group included in each Analysis  PLEASE PROVIDE NUMBERS (in absolute numbers rather than percentages)
		Was the analysis by intention to treat?
17	Results: Outcomes and estimation	For each primary and secondary outcome are the following included?
		Summary of results
		Estimated effect size
		Precision (eg. 95% confidence interval)
	Please state the results for each outcome measure	
18	Results: Ancillary analyses	Address multiplicity by reporting any other analyses performed, including sub group and adjusted analyses, indicating whether those preplanned, and those exploratory
19	Results: Adverse events	All important adverse events or side effects in each Group
20	Discussion: Interpretation	Interpretation of the results, taking into account study hypotheses, sources of potential bias or imprecision, and the dangers associated with multiplicity of analyses outcomes
21	Discussion: Generalizability	Extent of generalisability
22	Discussion: Overall evidence	General interpretation of results in the context of current evidence

<sup>1</sup>Altman, D. G., Schulz, K. F., Moher, D., Egger, M., Davidoff, F., Elbourne, D., Gøtzsche, P. C., Lang, T. (2001) The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Annals of Internal Medicine*, **134**, 663-694.

<sup>2</sup>Critical Appraisal Skills Programme (accessed 14/05/05) 10 questions to help you make sense of randomised clinical trials. <http://www.phru.nhs.uk/casp/rect.pdf>



# **APPENDIX IV. ASSESSING THE QUALITY OF RANDOMISED CLINICAL TRIAL REPORTS FOR A PROPOSED REVIEW: IS BLINDING NECESSARY?**

This assignment was written by Catherine Minns Lowe and submitted as part of the Systematic Reviews Module, Department of Continuing Education, University of Oxford 01/07/2005. It was awarded Grade A.

## **Introduction.**

This assignment addresses the question: is it necessary for systematic reviewers assessing the quality of trial reports in a proposed systematic review, evaluating the effectiveness of post-operative physiotherapy exercise following hospital discharge for patients undergoing elective total hip/knee arthroplasty, (Sub Appendix I), to be blinded to certain details of the report? And, if so, to what aspects of the trials should blinding be applied?

## **Background.**

The issue of blinding is not new. Nearly 20 years ago it was suggested that the quality of reports should be assessed under blind conditions to prevent/minimise the introduction of selection and data-extraction bias into the review process<sup>1</sup>. However, until recently, little research has been undertaken to assess whether blinding increases/reduces such bias<sup>2-4</sup>. Contradictory evidence supports<sup>4-6</sup> or refutes<sup>2,7-9</sup> whether blinding reviewers affects the quality or outcome of the review process and introduces bias. An ongoing systematic review may provide more definitive answers but has yet to be completed<sup>10</sup>.

## **How blinding might effect our review.**

Information was obtained (Sub Appendix II) and considered with regard to our review and its constraints and practical limitations.

Blinding reviewers would be problematic for our review. The review team consists of three physiotherapy researchers; with no funding for extra assistance. Reviewer 1 has identified studies and, through peer review with Reviewer 3, has selected studies for inclusion into the review; both cannot be blinded for the quality reviewing stage. In addition, Reviewer 1 has in-depth knowledge of the review topic and, even if blinding could have taken place, would be highly likely to recognise studies (a recognised difficulty/limitation of the blinding process<sup>2</sup>) and may be less critical of their quality because of her interest and enthusiasm for the topic<sup>11</sup>. Such prior knowledge of study results may be a problem since studies may be selected/rejected according to the reviewer's preconceived ideas about the topic<sup>10</sup>. Reviewer 1 believes she has no preference regarding the outcome of the review but may be unconsciously biased<sup>12</sup>. Reviewer 2 was not involved in study selection and is able to be blinded for the data extraction stage to rate study quality. Reviewer 2 has a background in physiotherapy research but no specific knowledge of joint replacement physiotherapy research and is unlikely to have previously read the studies. She may therefore be more critical of the quality of studies because of lower professional interest in the topic<sup>11</sup> and, although as an experienced researcher blinding may be less

successful<sup>7</sup>, she may review more strictly<sup>5</sup>. For these reasons it would be inappropriate to randomise blinding / non-blinding status for the reviewers<sup>10</sup> but we could allocate a blind and non blind reviewer. Other research indicates that reviewer characteristics have little effect on reviewing quality<sup>13</sup> so a variety of views exists. Surprisingly little is known about the cognitive aspects of reviewing<sup>12</sup> and other factors may influence how our reviewers rate study quality that are not yet predictable.

If we include blinding we need to decide what study aspects to blind and, again, considerable variety occurs. Blinding for author<sup>1-2,4,7-10</sup>, because either well-known authors may be rated less critically<sup>2</sup> or more harshly due to professional jealousies<sup>5</sup>, is unlikely to be relevant to review since there are no renowned experts in our review topic area. The identified research is by different authors with no programmes of research in the area. However, there is the risk of reviewers personally knowing authors. Blinding for affiliation<sup>2,4,7,8,10</sup>, acknowledgements<sup>4</sup>, funding source<sup>4</sup> or location<sup>14</sup> appears relevant since recognised centres of excellence for physiotherapy research exist and may influence reviewer's perception of quality. Blinding for journal<sup>1,4,10,14</sup>, and the look of the journal<sup>10</sup>, may also be important since our reviewers may give greater import to high repute journals. Blinding for study results<sup>10</sup> and identity of control and treatment groups<sup>3</sup> might also influence our review if, as mentioned previously, preconceptions about treatment effect are held by reviewers: however, this is not thought to be true.

Blinding methods<sup>10</sup> are another issue for our review. Concealment ranges from electronically scanning and editing reviews<sup>3</sup> to using a black marker/opaque tape to score out relevant aspects<sup>6,14</sup>. We have no scanner so the latter would be a practical and less time consuming<sup>3,6</sup> choice. The extent of blinding also varies<sup>10</sup>; some researchers remove relevant aspects throughout the study<sup>3</sup>, others only from the title page, headers/footers and acknowledgements and not from the main text where identifiers may still be included<sup>2,7,8-9</sup>. Elsewhere this is unclear<sup>4,14</sup>.

Advocates of blinding reviewers believe this may produce less biased reviews<sup>5</sup> of higher quality<sup>6</sup> and with lower and more consistent scores than non blinded ones<sup>4</sup> whilst other research suggests that blinding is not achievable or worth the time and effort required<sup>2,4-5,7,9</sup>

For example, Berlin (1997<sup>3</sup>) concluded blinding was not worthwhile. Berlin's research was then criticised for the small number of meta-analyses, plus the considerable heterogeneity, involved<sup>15</sup>. The low rates of blinding achieved, ranging from 54-73%<sup>2,5-6,9</sup> are another reason why some advocate blinding to be inappropriate/unnecessary. Additionally, some reviewers find blind reviewing more difficult<sup>3</sup>; as our reviewers are experienced this is unlikely to be a problem.

### **Decision.**

Using the literature alone, no definitive answer was discovered although there seems to be more support for the view that blinding is not successful or worthwhile. However, since debate exists and conclusive evidence is absent, decisions regarding blinding rest with the authors and the available resources<sup>16</sup>. For our review Reviewer 1 will not be blinded to the review whilst Reviewer 2 will be blinded. Success/failure rates of blinding will be assessed and interobserver agreement will be measured as advised<sup>1</sup>. This approach of using blinded and non blinded reviewers and comparing the results has previously been used successfully<sup>9</sup>. Reviewer 1 will conceal affiliations, acknowledgements, funding, location and journal throughout the documents for Reviewer 2, trying to avoid creating incomprehensible text<sup>10</sup>.

Following independent review by Reviewers 1 and 2, discrepancies will be discussed in an attempt to reach consensus<sup>17</sup>. Discrepancies will also be discussed with Reviewer 3 to assist achieving consensus. We will clearly state the processes we have used in our review write up<sup>16</sup> and believe that having a blinded reviewer will assist acceptance for publication in addition to contributing to the quality of our review. Additionally, we accept that blinding is only one aspect of a review's quality and that all aspects of the review effecting quality, such as the quality review instrument used<sup>9</sup> need to be adequately addressed.

## REFERENCES.

1. Sacks, H. S., Berrier, J., Reitman, D., Acona-Berk, V. A., Chalmers, T. C. (1987) Meta-analysis of randomised control trials. *New England Journal of Medicine* **316**(8): 450-455.
2. Justice, A. C., Cho, M. K., Winker, M. A., Berlin, J. A., Rennie, D., and the PEER investigators (1998) Does masking author identity improve peer review quality? A randomized controlled trial. *Journal of the American Medical Association* **280**(3): 240-242.
3. Berlin, J. A. (1997) Does blinding of readers affect the results of meta-analyses? *The Lancet* **350**: 185-186.
4. Jadad, A. R., Moore, A., Carroll, D., Jenkinson, C., Reynolds, D. J. M., Gavaghan, D. J., McQuay, H. J. (1996) Assessing the quality of reports of randomised clinical trials: is blinding necessary? *Controlled Clinical Trials* **17**: 1-12.
5. Fisher, M., Friedman, S. B., Strauss, B. (1994) The effects of blinding on acceptance of research papers by peer review. *Journal of the American Medical Association* **272**(2): 143-146.
6. McNutt, R. A., Evans, A. T., Fletcher, R. H., Fletcher, S. W. (1990) The effects of blinding on the quality of peer review. *Journal of the American Medical Association* **263**(10): 1371-1376.
7. Cho, M. K., Justice, A. C., Winker, M. A., Berlin, J. A., Waeckerle, J. F., Callahan, M. L., Rennie, D., and the PEER investigators (1998) Masking author identity in peer review. What factors influence masking success? *Journal of the American Medical Association* **280**(3): 243-245.
8. Godlee, F., Gale, C. R., Martyn, C. N. (1998) Effect on the quality of peer review of blinding reviewers and asking them to sign their reports. *Journal of the American Medical Association* **280**(3): 237-240.
9. Van Rooyen, S., Godlee, F., Evans, S., Smith, R., Black, N. (1998) Effect of blinding and unmasking on the quality of peer review. *Journal of the American Medical Association* **280**(3): 234-237.
10. Westby, M., Clarke, M., Hopewell, S., Ram, F. (2004) Masking reviewers at the study inclusion stage in a systematic review of health care interventions. (Protocol).

*The Cochrane Database of Methodology Reviews* Issue 2. Art. No.:MR000019. DOI: 10.1002/14651858. MR000019.

11. Glantz, S. A., Bero, L. A. (1994) Inappropriate and appropriate selection of 'peers' in grant review. *Journal of the American Medical Association* **272**(2): 114-116.
12. Kassirer, J. P., Campion, E. W. (1994) Peer review: crude and understudied, but indispensable. *Journal of the American Medical Association* **272**(2): 96-97.
13. Black, N., van Rooyen, S., Godlee, F., Smith, R. Evans, S. (1998) What makes a good reviewer and a good review for a general medical journal? *Journal of the American Medical Association* **280**(3): 231-233.
14. Moher, D., Fortin, P., Jadad, A. R., Jüni, P., Klassen, T., Le Lorier, J., Liberati, A., Linde, K., Penna, A. (1996) Completeness of reporting of trials published in languages other than English: implication for conduct and reporting of systematic reviews. *Lancet* **347**: 363-366.
15. West, R. (1997) Does blinding of readers affect the results of meta-analysis? *Lancet* **350**: 892.
16. Irwig, I. Tosteson, A. N. A., Gatsonis, C., Lau, J., Colditz, G., Chalmers, T. C., Mosteller, F. (1994) Guidelines for meta-analyses evaluating diagnostic tests. *Annals of Internal Medicine* **120**(8): 667-676.
17. L'Abbe, K. A., Detsky, A. S., O'Rourke, K. (1987) Meta-analysis in clinical research. *Annals of Internal Medicine* **107**(2): 224-33.

## BIBLIOGRAPHY.

- Callaham, M. L., Baxt, W. G., Waeckerle, J. F., Wears, R. L. (1998) Reliability of editor's subjective quality ratings of peer reviews of manuscripts. *Journal of the American Medical Association* **280**(3): 229-231.
- Chalmers, T. C., Celano, P., Sacks, H. S., Smith, H. (1983) Bias in treatment assignment in controlled clinical trials. *New England Journal of Medicine* **309**(22): 1358-1361.
- Cochrane Reviewer's Handbook 4.4.2. Chapter 6.9. Application of Quality Qssessment. <http://www.cochrance.dk/cochrane/handbook/hbook.htm>
- Egger, M., Davey Smith, G., Altman, D. G., editors (2001) *Systematic Reviews in Health Care: Meta-analysis in Context*, 2<sup>nd</sup> edition. London: BMJ Publishing Group.
- University of Oxford Department for Continuing Education: Continuing Professional Development Centre (2005) *Systematic Reviews* course notes. Course Tutor: Mike Clarke. Oxford: Continuing Professional Development Centre.

## ADDITIONAL.

The following paper was identified and ordered but did not arrive in time for inclusion into this assignment.

Berlin, J. A., Miles, C. G., Crigliano, M. D. *et al.*, (1997) Does blinding of readers affect the results of meta-analyses? Results of a randomized trial. *Online J Curr Clin Trials*. Document no 205.

## *SUB APPENDIX I.*

### A SUMMARY OF OUR PROPOSED REVIEW.

**Title:** A systematic review evaluating the effectiveness of post operative physiotherapy exercise following discharge from hospital (in terms of improving function, mobility and strength) for patients undergoing elective total hip or knee arthroplasty (joint replacement).

**Background:** Generally there is a marked and rapidly decreasing length of stay following arthroplasty. Previous work suggests that early discharge after arthroplasty frequently results in persistent muscle weakness and poor functional mobility and that further rehabilitation within primary care may be required to optimise recovery for these patients. It is not yet known whether such physiotherapy interventions are effective in improving function, mobility, strength or range of movement.

**Purpose:** This review is primarily being undertaken to inform the development of a new complex intervention for use in a randomized controlled trial “Is post discharge physiotherapy effective in improving function and mobility following elective joint arthroplasty in a Diagnostic and Treatment Centre?” The trial is due to start in October 2005. The purpose of the trial is to discover whether patients discharged early following elective joint arthroplasty significantly benefit from additional physiotherapy post discharge, compared to usual care.

**Participants:** The participants for inclusion in this review are patients undergoing elective total hip or knee arthroplasty who have received physiotherapy following discharge from hospital post-operatively.

**Intervention:** The intervention “physiotherapy” is referred to by different names in the literature; namely, physiotherapy, physical therapy, home programme, exercise and rehabilitation. For the purposes of this review the term physiotherapy refers to any exercises or exercise programme provided by physiotherapists/physical therapists during the rehabilitative period after discharge from hospital after joint replacement surgery occurring in the out patient, community or home setting.

**Comparison:** The review will include randomized clinical trials relevant to the review question, regardless of language of publication. Trials will be included if they

compare a post discharge physiotherapy exercise intervention compared with either control group, usual care or another form of physiotherapy exercise intervention.

**Outcomes:** The relevant outcomes by which effectiveness is to be measured include the following: Function, Mobility, Muscle strength, Range of motion.

**Location of Studies:** Early searches were conducted prior to this systematic review to begin to identify appropriate search terms. Since these searches indicated that a large number of terms seemed relevant, a broad area search was initially undertaken which was followed by a number of more specific searches to check that as many relevant records as possible were identified. It was accepted that this inclusive approach would identify many papers irrelevant to the review but this approach seemed to offer the most comprehensive method of locating the maximum possible number of trials present. The databases searched were: AMED 1985-; CINAHL 1982-; EMBASE 1974-; KingsFund 1979-; MEDLINE 1966-; The Cochrane library (including Cochrane reviews, Cochrane Central Register of Controlled Trials, DARE); PEDro physiotherapy evidence database; The Department of Health National Research Register

**Terms searched** included:

Hip, knee, replacement, arthroplasty, trial, exercise, physiotherapy and physical therapy, home programme.

**Selection of Studies:** Two reviewers will independently assess the eligibility of identified studies for inclusion into the review against specified criteria. A consensus meeting will be held to discuss discrepancies.

**Data Extraction:** Data regarding the quality of the studies and their results will be independently extracted by two reviewers. A consensus meeting will be held to discuss discrepancies and a third reviewer is available for further discussion.

**Quality:** A form, based upon the CONSORT statement<sup>1</sup> and the CASP guidelines<sup>2</sup> has been developed to extract information for the review. This form includes rating the following aspects: adequate randomisation procedures; adequate concealment; blinding of outcome assessment; attrition; power and size of study; effect size.

**Analysis:** This stage of the protocol is still being developed. A statistician is being consulted regarding the design of this section. It is anticipated that, since the studies included in this view are not likely to be heterogenous, a meta-analysis will be inappropriate and potentially misleading for this review.

1. Altman, D. G., Schulz, K. F., Moher, D., Egger, M., Davidoff, F., Elbourne, D., Gøtzsche, P. C., Lang, T. (2001) The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Annals of Internal Medicine*, **134**: 663-694.

2. Critical Appraisal Skills Programme (accessed 14/05/05) 10 questions to help you make sense of randomised clinical trials. <http://www.phru.nhs.uk/casp/rct.pdf>

***SUB APPENDIX II.***

To identify information for this assignment on line databases, including Medline, Embase and Cinahl, were searched using the title term “blinding” AND document term “systematic review”. The Cochrane Library Methodology Reviews, Cochrane Reviewer’s Handbook 4.4.2 and the National Coordinating Centre for Health Technology Assessment Monograph Series on CD-ROM were also searched. Relevant papers identified from these searches were obtained and their reference lists and citations were searched for further publications, as were references lists from course notes and the Systematic Reviews Course’s recommended text (see bibliography). This search strategy was not intended to be 100% exhaustive but was designed to provide sufficient relevant and appropriate information for this 1000 word assignment.

## APPENDIX V. Blinding Monitoring Form.

1. Can you identify the authors of this research report? *(Please tick one response)*

Yes  No  Think I can guess

If you have answered 'YES' or 'Think I can guess', then please write the author/s below

.....  
.....

2. Can you identify any author affiliations of this research report? *(Please tick one response)*

Yes  No  Think I can guess

If you have answered 'YES' or 'Think I can guess', then please write the affiliations below

.....  
.....

3. Can you identify the funding source for this research? *(Please tick one response)*

Yes  No  Think I can guess

If you have answered 'YES' or 'Think I can guess', then please write the funding source below

.....  
.....

4. Can you identify the location/s at which this research took place? *(Please tick one response)*

Yes  No  Think I can guess

If you have answered 'YES' or 'Think I can guess', then please write the location/s below

.....  
.....

5. Can you identify the journal this research report is published in? *(Please tick one response)*

Yes  No  Think I can guess

If you have answered YES, then please write the Journal name below



.....  
.....

**APPENDIX VI. DATA EXTRACTION FORM FOR SYSTEMATIC REVIEW** (based on the CONSORT1 statement and CASP2 guidelines.)

Authors and Year of Publication:

Item Number	Paper Section and Topic	Descriptor: Does the study include/ mention this?	YES or NO or Partial or Unclear	Appropriate / Inappropriate	Comments / Details
1	Title & abstract	how participants were allocated to interventions? (eg: “random allocation” “randomized”)			
2	Introduction: Background	the scientific background and explanation of rationale included?			
3	Methods: Participants	Eligibility criteria			
		Method of recruitment			
		Setting and locations where data was collected			
4	Methods: Interventions	Precise details of the interventions intended for each group and how and when they were actually administered			
5	Methods: Objectives	Specified objectives and hypotheses			

6	Methods: Outcomes	Clearly defined primary and secondary outcome Measures.			
		Please state outcome measures used			
		When applicable, any methods used to enhance the quality of measurement (e.g. multiple observation, training of assessors)			
7	Methods: Sample size	How sample size was determined and, when applicable, explanation of any interim analyses  PLEASE STATE SAMPLE SIZE			
	Methods: randomisation	Does it state that randomisation has occurred?			
8	Methods: Randomisation Sequence generation	Method used to generate the random allocation sequence, including details of any restriction (Eg: blocking or stratification)			
9	Methods: Randomisation Allocation concealment	Method used to implement the random allocation Sequence, cf: whether the sequence was concealed until interventions were assigned			

<b>10</b>	Methods: Randomisation Implementation	Who generated the allocation sequence, who enrolled Patients, and who assigned patients to their groups			
<b>11</b>	Methods: Blinding	Were participants blinded to allocation? (Consider that this may not be possible <sup>2</sup> )			
		Were those administering the intervention blinded? (Consider that this may not be possible <sup>2</sup> )			
		Was outcome assessment blinded?			
<b>12</b>	Methods: Statistical methods	Methods used to compare groups for primary outcomes and methods for additional analyses (such as subgroup and adjusted). Please state.			
<b>13</b>	Results: Participant flow	Flow of participants through each stage (i.e. number Randomized, receiving intended treatment, completing protocol, analysed. Also, protocol deviations (& reasons) as planned.			
<b>14</b>	Results: Recruitment	Dates defining the periods of recruitment and follow-up			
<b>15</b>	Results: Baseline data	Baseline demographics and clinical characteristics presented for each group			
<b>16</b>	Results: Numbers analysed	No of participants in each group included in each Analysis  PLEASE PROVIDE NUMBERS (in absolute numbers rather than percentages)			
		Was the analysis by intention to treat?			
<b>17</b>	Results: Outcomes and estimation	For each primary and secondary outcome are the following included?			
		Summary of results			

		Estimated effect size			
		Precision (eg. 95% confidence interval)			
	Please state the results for each outcome measure				
<b>18</b>	Results: Ancillary analyses	Address multiplicity by reporting any other analyses performed, including sub group and adjusted analyses, indicating whether those preplanned, and those exploratory			
<b>19</b>	Results: Adverse events	All important adverse events or side effects in each Group			
<b>20</b>	Discussion: Interpretation	Interpretation of the results, taking into account study hypotheses, sources of potential bias or imprecision, and the dangers associated with multiplicity of analyses			

		outcomes			
21	Discussion: Generalizability	Extent of generalisability			
22	Discussion: Overall evidence	General interpretation of results in the context of current evidence			

<sup>1</sup>Altman, D. G., Schulz, K. F., Moher, D., Egger, M., Davidoff, F., Elbourne, D., Gøtzsche, P. C., Lang, T. (2001) The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Annals of Internal Medicine*, **134**, 663-694.

<sup>2</sup>Critical Appraisal Skills Programme (accessed 14/05/05) 10 questions to help you make sense of randomised clinical trials. <http://www.phru.nhs.uk/casp/rct.pdf>

## APPENDIX VII. LITERATURE SEARCH STRATEGIES.

Initially a broad area search was undertaken which was followed by a number of more specific searches to check that as many relevant records as possible were identified.

Terms searched included:

Hip, knee, replacement, arthroplasty, trial, exercise, physiotherapy and physical therapy. The full search strategies are listed below.

### SEARCH 1.

**Date:** 10-11<sup>th</sup> March 2005

**Databases:** Searched the following clinical databases via KA 24

AMED 1985-  
CINAHL 1982-  
EMBASE 1974-  
KingsFund 1979-  
MEDLINE 1966-

#### Search terms:

“hip” OR “knee”  
AND “replacement” OR “arthroplast\$”  
AND “rehabilitation” AND “trial\$”

The whole documents were searched.

**Hits:** Total number of hits = 747

After removal of duplicates command the number of hits = 595

After hand searching the records more duplicates were found and 8 further records were removed. The final numbers of hits = 587

**Results:** The records were searched and the table below shows the subject areas found.

Subject	Number
Records relating to rehabilitation and / or exercise	56
	Of these 25 abstracts related to post operative exercise / rehabilitation 11 related to pre operative exercise therapeutic modalities (e.g. TNS, electrical stimulation) = 6 gait = 5 muscle training = 4 weight bearing = 1 manipulation = 1 overview = 1 In patient exercise=1 Home programme = 1
Records relating to other conditions (for example, fractured neck of femurs, cruciate ligaments)	317
Records relating to operative and postoperative procedures, complications and drugs.	50
Records relating to the condition Osteoarthritis	42
Records relating to predictors of outcome or outcome measures	25
Records relating to patient information &/or education	20

Records relating to the use of continuous passive motion	22
Records providing reviews or opinions	17
Records relating to prostheses	15
Records relating to pain	7
'other' records (relating to aspects like nutrition, fatigue, practice guidelines, patient / carer experience, cost effectiveness of hospital at home schemes, length of stay)	33

These numbers total 604 because some records belonged in more than one category.

**Updated 24/04/2007:** 180 hits  
Nil new records.

## SEARCH 2

Following the above search, several more specific searches were undertaken to enhance likelihood of all possible relevant records being identified.

**Date:** 16<sup>th</sup> March 2005

**Databases:** Searched the following clinical databases via KA 24  
 AMED 1985-  
 CINAHL 1982-  
 EMBASE 1974-  
 KingsFund 1979-  
 MEDLINE 1966-

### Search terms:

"hip" OR "knee" (whole document)  
 AND "replacement" OR "arthroplast\$" (whole document)  
 AND "exercise" AND "trial\$" (title)

**Hits:** Total number of hits = 198  
 After removal of duplicates command the number of hits = 118

3 new relevant records, on post operative exercise were identified for consideration for inclusion into the review. A further 8 records, relevant to the proposed trial, were located. Of these, 3 were review articles, one record related to recreational exercise after joint arthroplasty, 3 to pre operative exercise, and one to inpatient pre and post operative exercise.

**Updated 24/04/2007:** 1 hits  
Nil new records.

## SEARCH 3

**Date:** 16<sup>th</sup> March 2005

**Databases:** Searched the following clinical databases via KA 24  
 AMED 1985-  
 CINAHL 1982-  
 EMBASE 1974-  
 KingsFund 1979-  
 MEDLINE 1966-

### Search terms:

"hip" OR "knee" (whole document)  
 AND "replacement" OR "arthroplast\$" (whole document)  
 AND "physiotherapy" AND "trial\$" (title)



**Hits:** Total number of hits = 6  
After removal of duplicates command the number of hits = 2

**Results:** No new relevant records found.

**Updated 24/04/2007:** 4 hits  
Nil new records.

#### **SEARCH 4**

**Date:** 16<sup>th</sup> March 2005

**Databases:** Searched the following clinical databases via KA 24

AMED 1985-  
CINAHL 1982-  
EMBASE 1974-  
KingsFund 1979-  
MEDLINE 1966-

**Search terms:**

“hip” OR “knee” (whole document)  
AND “replacement” OR “arthroplast\$” (whole document)  
AND physiotherapy (title)

**Hits:** Total number of hits = 63  
After removal of duplicates command the number of hits = 39

**Results:** No new records for consideration for inclusion into the review were identified. New related records identified = 5. One record related to a post operative complication, one to a conference paper about preoperative physiotherapy, one to post operative physiotherapy and 2 were reviews.

**Updated 24/04/2007:** 14 hits  
Nil new records.

#### **SEARCH 5**

**Date:** 16<sup>th</sup> March 2005

**Databases:** Searched the following clinical databases via KA 24

AMED 1985-  
CINAHL 1982-  
EMBASE 1974-  
KingsFund 1979-  
MEDLINE 1966-

**Search terms:**

“hip” OR “knee” (whole document)  
AND “replacement” OR “arthroplast\$” (whole document)  
AND physical therapy (title)

**Hits:** Total number of hits = 66  
After removal of duplicates command the number of hits = 43

**Results:** 8 new records were identified for consideration for inclusion in to the review. A further 13 new relevant related records identified were identified for the proposed trial: 7 reviews or opinions, 3 comparing continuous passive motion and physical therapy, 1 relating to prosthetic type, 1 ‘other’ relating to service utilisation and 1 relating to osteoarthritis.

**Updated 24/04/2007:** 15 hits  
Nil new records.

#### **SEARCH 6.**

**Date:** 16<sup>th</sup> March 2005

**Database:** The Cochrane library (including Cochrane reviews, Cochrane Central Register of Controlled Trials, DARE)

#### **Search:**

Browsed by topic musculoskeletal  
Search narrowed osteoarthritis 37 hits  
Search narrowed rehabilitation 9 hits  
Search narrowed exercise 3 hits

**Results:** 1. Intensity of exercise for the treatment of osteoarthritis. 2. Home versus centre based activity programmes for older adults. 3. Exercise for osteoarthritis of the knee

Patient education 1 hit

Continuous passive motion 1 hit

Modalities 1 (electromagnetic fields)

Splints / orthoses 1 hit

Multidisciplinary rehabilitation 1 hit

**Result:** protocol for Multidisciplinary rehabilitation interventions for joint replacement at the knee and hip for arthropathies.

To check the search, a further search of the term “joint replacement” was done. Of the 80 hits identified, there were no other relevant records.

**Updated** 24/04/2007: Nil new relevant records.

#### **SEARCH 7.**

**Date:** 16<sup>th</sup> March 2005

**Database:** PEDro physiotherapy evidence database

#### **Search terms:**

joint replacement AND rehabilitation

**Result:** 1 hit (record already identified in search 1)

**Search terms:** The search terms were then broadened to,  
joint replacement

**Results:** 5 hits. 1 relevant record (already identified in search 1) therefore no new records were identified via this search.

**Updated** 24/04/2007: Nil new relevant records.

#### **SEARCH 8.**

In addition to searching for published research, efforts were made to identify ongoing relevant research projects.

**Date:** 16<sup>th</sup> March 2005

**Database:** the Department of Health National Research Register.

**Search terms:** joint replacement AND rehabilitation .

**Result:** 0 hits

**Search terms:** joint replacement AND physiotherapy

**Results:** 7 hits - 2 ongoing, 5 completed. Ongoing projects included a pilot project of class versus outpatient physiotherapy after knee replacement, plus a project determining prognostic factors for joint replacement

outcomes. Completed projects included 1 on osteoarthritis, 1 on rheumatoid arthritis, 1 on acupuncture pre operatively, 1 on health beliefs and 1 on patient satisfaction.

**Search terms:** Joint replacement AND exercise

**Results:** 2 hits, both already identified in the previous search

**Search terms:** Joint replacement AND Physical Therapy

**Results:** 3 hits, no new records

**Search terms:** joint arthroplasty AND physiotherapy

**Results:** 0 hits

**Search terms:** joint arthroplasty AND rehabilitation

**Results:** 2 hits, one relevant new record relating to pre operative advice.

**Updated 24/04/2007:** 1 new trial protocol record. Lead investigator contacted. Recruitment failed therefore trial halted and not completed.

#### **SEARCH 9.**

Following reading of abstracts and initial reading of papers, a further search including “home programme” was performed since this one a further way by which exercise / rehabilitative approaches were termed.

**Date:** 24<sup>th</sup> March 2005

**Database:** Searched the following clinical databases via KA 24

AMED 1985-  
CINAHL 1982-  
EMBASE 1974-  
KingsFund 1979-  
MEDLINE 1966-

**Search terms:**

“hip” OR “knee” (whole document)  
AND “replacement” OR “arthroplast\$” (whole document)  
AND “home programme”(title)

**Hits:** Total number of hits = 7

After removal of duplicates command the number of hits =2

**Result:** Both trials had already been found in search 1.

**Updated 24/04/2007:** 1 hits  
Nil new records.

#### **SEARCH 10.**

Since so few articles were found in search 9, the search was then broadened

**Date:** 24<sup>th</sup> March 2005

**Database:** Searched the following clinical databases via KA 24

AMED 1985-  
CINAHL 1982-  
EMBASE 1974-  
KingsFund 1979-  
MEDLINE 1966-

**Search terms:**

“hip” OR “knee” (whole document)  
AND “replacement” OR “arthroplast\$” (whole document)  
AND “home programme”(whole document)

**Hits:** Total number of hits = 31

After removal of duplicates command the number of hits =22

**Result:** No further trials were identified. One new relevant review, one article on patient's perspective of joint replacement provision and one descriptive study were found; these were considered relevant to the proposed trial but not for the systematic review.

**Updated 24/04/2007:** 27 hits  
Nil new records.

#### **SEARCH 11.**

Hand searches of key journal and conference proceedings occurred between June 27<sup>th</sup> and July 6<sup>th</sup> 2005. The supervisors and applicant discussed and decided which sources would be most appropriate to search. The following were searched:

*Physiotherapy* Journal 1985-2004 inclusive. Contents page of each issue searched. No new relevant articles identified.

*Physical Therapy* 1985-2004 inclusive. Contents pages of each issue searched. 12 issues were missing from the library:

Vol 65 July 1985

Vol 65 October 1985

Vol 65 November 1985

Vol 66 September 1986

Vol 69 July 1989

Vol 69 October 1989

Vol 70 May 1990

Vol 71 August 1991

Vol 73 February 1993

Vol 73 June 1993

Vol 74 June 1994

Vol 74 September 1994

For years with missing journals, the annual index to the journal was fully searched.

No new relevant articles were identified for the systematic review.

#### *Journal of Bone and Joint Surgery (Britain) Conference Proceedings*

1985-2004 inclusive. The Proceedings and Reports of Universities, Colleges, Associations and Societies were searched until 1991. From 1992 onwards this section was replaced by the Orthopaedic Proceedings Supplements, which were searched.

**Results.** 2 relevant new abstracts were identified for consideration for inclusion into the systematic review. One, from 2004 does not appear to be published yet (author written to requesting further information). The second abstract was then looked up via KA24 (on a Medline, CINAHL and EMBASE search) and details of a full paper located and obtained.

**Updated 21/04/2007:** 1 new record in *JBJS*, duplicate of trial already included.

#### **SEARCH 12.**

One supervisor also suggested searching for relevant articles from the Occupational Therapy evidence base since rehabilitation following joint replacement is also a part of occupational therapy treatment..

**Date:** 6<sup>th</sup> July 2005

**Databases:** Searched the following clinical databases via KA 24

AMED 1985-

CINAHL 1982-

EMBASE 1974-  
KingsFund 1979-  
MEDLINE 1966-

**Search terms:**

“hip” OR “knee”

AND “replacement” OR “arthroplast\$”

AND “occupational therapy ” AND “trial\$”

The whole documents were searched.

**Hits:** Total number of hits = 0

**Updated 24/04/2007:** 0 hits  
Nil new records.

**SEARCH 13.**

**Date:** 6<sup>th</sup> July 2005

**Databases:** Searched the following clinical databases via KA 24

AMED 1985-  
CINAHL 1982-  
EMBASE 1974-  
KingsFund 1979-  
MEDLINE 1966-

**Search terms:**

“hip” OR “knee” (whole document)

AND “occupational therapy” (title)

**Hits:** Total number of hits = 55

After removal of duplicates command the number of hits = 35

**Results:** No new records identified.

**Updated 24/04/2007:** 3 hits  
Nil new records.

**SEARCH 14.**

**Date:** 6<sup>th</sup> July 2005

**Databases:** Searched the following clinical databases via KA 24

AMED 1985-  
CINAHL 1982-  
EMBASE 1974-  
KingsFund 1979-  
MEDLINE 1966-

**Search terms:**

“hip” OR “knee” (whole document)

AND “occupational therapist\$” (title)

**Hits:** Total number of hits = 7

**Results:** No new records identified.

**Updated 24/04/2007:** 3 hits  
Nil new records.

**SEARCH 15.**

**Date:** 6<sup>th</sup> July 2005

**Database: Database:** the Department of Health National Research Register.

**Search terms:** joint replacement AND occupational therapy.

**Result:** 5 hits. No new relevant records identified for consideration for the review.

**Updated** 24/04/2007: Nil new records.

## APPENDIX VIII. REASONS FOR THE SELECTION OF PAPERS INCLUDED IN /EXCLUDED FROM THE REVIEW.

### A. Full papers/details obtained and considered for inclusion/exclusion.

Codine *et al.*, 2004; trial to evaluate the use of low velocity submaximal eccentric hamstring contractions on recovery of extension (day 10-30 post op). INCLUDED in the review.

Frost *et al.*, 2002; pilot trial after TKR functional versus traditional exercise. INCLUDED in the review.

Grange *et al.*, 2004; restricted rct (n=14) of arm crank intervention post operatively following THR. Unclear from abstract when the intervention occurred. (?when post op). Subsequently EXCLUDED from the review (19<sup>th</sup> May 2005) after the full paper was read since the intervention took place whilst the participants were inpatients at a rehabilitation centre.

Gursen and Ahrens, 2004; the full paper was obtained since there was no available abstract. Upon reading the full paper, it was EXCLUDED from the review (14<sup>th</sup> July 2005) since the trial intervention was a pre operative home visit.

Hauer *et al.*, 2002; trial to evaluate a training after falls rehabilitation intervention. The intervention was provided 6-8 weeks post hip surgery (following fracture or elective hip TJR). The population was a sub population of a larger study of patients with a history of injurious falls. The paper was subsequently EXCLUDED (19<sup>th</sup> May 2005) since 25 / 28 patients were hip fractures and only 3 were elective joint replacements and the surgery was mixed (TJR = 14; hemiarthroplasty = 4; osteosynthesis = 7).

Hesse *et al.*, 2003; trial of treadmill training versus control following total knee joint replacement. It was not clear from abstract when the intervention occurred. The paper was subsequently EXCLUDED from the review (14<sup>th</sup> July, 2005) since on reading the full it seemed that the intervention occurred whilst the participants were inpatients.

Jan *et al.*, 2004; rct home programme following joint replacement. INCLUDED in the review.

Johnsson *et al.*, 1988; trial of Physiotherapy intervention 2/12 after THR. INCLUDED in the review.

Kramer *et al.*, 2003; trial of clinic and home based rehabilitation following TKR. INCLUDED in the review.

Maire *et al.*, 2004. trial of arm-interval training versus control following THR. From the abstract it was unclear when the intervention occurred. The full paper was subsequently EXCLUDED (19<sup>th</sup> May 2005) from the review since the intervention took place whilst participants were in a rehabilitation centre.

Mockland and Beverland, 2004; Abstract only for a trial evaluating whether outpatient physiotherapy improves range of motion after knee replacement, versus control. From the abstract we would INCLUDE this trial in the review therefore the authors were written too and further information requested and later full details were obtained. This enabled the study to be included.

Moffet *et al.*, 2004; intensive rehabilitation 2 months after discharge. INCLUDED in the review.

Patterson *et al.*, 1995; methodology unclear from abstract. Comparison of exercise program after THR with control. INCLUDED in the review.

Rajan *et al.*, 2004; rct comparing outpatient PT. INCLUDED in the review.

Suetta, *et al.*, 2004; rct comparing standard, muscle strengthening and EMS post THR. INCLUDED in the review.

Sashika *et al* 1996; Clinical trial THR – home program. INCLUDED in the review.

Trudelle Jackson *et al.*, 2004. rct of late ex programme after THR. INCLUDED in the review.

Worland *et al.*, 1998; trial comparing home Continuous Passive Motion machine versus a physiotherapy programme following total knee replacement. EXCLUDED from the review (19<sup>th</sup> May 2005) since CPM is not considered active exercise.

#### **B. Non English Language Papers. Papers reviewed.**

Nyberg and Kreuter, 2002. trial comparing group physical therapy and individual training at home in patients following THR (n=33). This Swedish paper was translated by a professional translation service for INCLUSION in the review (10<sup>th</sup> June 2006).

Lapshin *et al.*, 2002. Russian paper re: therapeutic ex in rehabilitation of elderly and aged patients after hip joint replacement. No English abstract available therefore the paper was obtained for translation. The paper was read by a colleague fluent in Russian who confirmed that the research was not a randomised clinical trial and the paper was therefore EXCLUDED (10<sup>th</sup> June 2006).

#### **C. Non English Language Papers. Abstracts reviewed.**

Chen *et al.*, 2004. Chinese paper. Abstract in English is difficult to fully understand. Rct (n=45) of routine and recovery rehabilitation groups following hip replacement. EXCLUDED from the review (19<sup>th</sup> May 2005) because the intervention appeared to be an inpatient intervention aimed at reducing post operative complications.

Kolarz *et al.*, 1999. German paper comparing young (n=58, aged <70) and older n=40 > age 70) patients receiving inpatient rehabilitation following knee replacement. EXCLUDED (19<sup>th</sup> May 2005) because the trial was for an inpatient intervention.

Nyberg and Kreuter, 2002. Swedish rct comparing group physical therapy and individual training at home in patients following THR (n=33). From the English abstract this paper was INCLUDED in the review and was translated by a professional translation service.

Tum-Sugden, 1976; The English abstract for this German paper does not describe a trial and includes ice in the treatment given to patients. The paper did not meet the inclusion criteria and was therefore EXCLUDED ((19<sup>th</sup> May 2005).

Ulreich *et al.*, 1997; German paper evaluating the effectiveness of a multidisciplinary rehabilitation programme after total knee replacement (n=65). EXCLUDED (19<sup>th</sup> May 2005) because this was a pilot study rather than a clinical trial and the intervention was an inpatient treatment.

Werner *et al.*, 2004; German paper comparing treadmill training with partial body weight support and traditional physiotherapy following THR. N = 80. From the abstract this appeared to be a duplicate of the study by Hesse *et al.*, 2003 (same authors, same results numbers) and was EXCLUDED from the review (19<sup>th</sup> May 2005).

#### **D. Excluded 19<sup>th</sup> May 2005 (Abstracts reviewed only).**

Beaupre *et al.*, 2001; intervention was post op inpatient regimes

Benedetti *et al.*, 2003; Longitudinal study to evaluate residual muscle function abnormalities after TKA.

Cullen *et al.*, 1973; Not a trial. Study is describing post THR rehabilitation n = 53.

Drabsch *et al.*, 1998; Descriptive study of task specific training on walking and sit-to-stand after total hip replacement.

Freburger, 2000; Not a trial. Study examining PT utilisation and outcomes (cost, discharge destination)



Gilbey *et al.*, 2003; intervention was both pre op (for 8 weeks) and post op (for 12 weeks)

Gilbey *et al.*, 2003 b. seems the same trial as above but in different journal.

Hewitt & Shakespeare, 2001; intervention was post op inpatient mobilisation regimes rather than post discharge.

Hughes *et al.*, 1993; Evaluated the effect of inpatient physiotherapy on hospital length of stay.

Jeudson and Stiller, 2002; intervention was post op inpatient regimes

Karst *et al.*, 1995; study evaluating in patient PT post op.

Kim and Moon, 1995; intervention was post op inpatient regimes

Kumar *et al.*, 1996; intervention was post op inpatient regimes

Lang, 1998; comparison of 6 & 7 day PT on length of hospital stay.

Licciardone *et al.*, 2004; trial for osteopathic manipulation following TJR rather than PT.

Montgomery & Eliasson, 1996; in patient trial only (CPM v Ex)

Munin *et al.*, 1998; rct of inpatient rehab only.

Richardson, 1975. Not a trial. Paper describes management of THR.

Ritter *et al.*, 1989; inpatient study re: CPM

Shih *et al.*, 1994; not a trial. A prospective study which aimed to quantitatively measure relative muscle torque strengths around the hip.

Stevens, J. E. 2002, rct evaluating neuromuscular stimulation after TKR.

Waters, 1974. Review of two different prosthetic TKRs.

## APPENDIX IX. STATISTICAL TESTS FOR REVIEWER AGREEMENT IN THE SYSTEMATIC REVIEW.

### KAPPA Crosstabs

Output Created	25-OCT-2006 13:37:55		
Comments			
Input	Data	E:\PhD\Systematic review\cmlsr2.sav	
	Filter	<none>	
	Weight	<none>	
	Split File	<none>	
	N of Rows in Working Data File	160	
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.	
	Cases Used	Statistics for each table are based on all the cases with valid data in the specified range(s) for all variables in each table.	
Syntax	CROSSTABS /TABLES=VAR00001 BY VAR00002 /FORMAT= AVALUE TABLES /STATISTIC=KAPPA /CELLS= COUNT /COUNT ROUND CELL .		
Resources	Elapsed Time	0:00:00.05	
	Dimensions Requested	2	
	Cells Available	116508	

#### Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
CML * KB	160	100.0%	0	.0%	160	100.0%

#### CML \* KB Crosstabulation

Count		KB				Total
		1.00	2.00	3.00	4.00	
CML	1.00	70	6	7	2	85
	2.00	4	36	2	5	47
	3.00	13	4	8	1	26
	4.00	1	1	0	0	2
Total		88	47	17	8	160

**Symmetric Measures**

	Value	Asymp. Std. Error(a)	Approx. T(b)	Approx. Sig.
Measure of Agreement Kappa	.524	.054	9.369	.000
N of Valid Cases	160			

a Not assuming the null hypothesis.

b Using the asymptotic standard error assuming the null hypothesis.

**Reliability ICC (2,1) Absolute Agreement**

Output Created	25-OCT-2006 14:19:02		
Comments			
Input	Data	E:\PhD\Systematic review\cmlsr1.sav	
	Filter	<none>	
	Weight	<none>	
	Split File	<none>	
	N of Rows in Working Data File	165	
	Matrix Input	E:\PhD\Systematic review\cmlsr1.sav	
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.	
	Cases Used	Statistics are based on all cases with valid data for all variables in the procedure.	
Syntax	RELIABILITY /VARIABLES=VAR00001 VAR00002 /FORMAT=NOLABELS  /SCALE(ALPHA)=ALL/MODEL=ALPHA /ICC=MODEL(RANDOM) TYPE(ABSOLUTE) CIN=95 TESTVAL=0 .		
Resources	Elapsed Time	0:00:00.03	
	Memory Available	524288 bytes	
	Largest Contiguous Area	524288 bytes	
	Workspace Required	128 bytes	

**Case Processing Summary**

		N	%
Cases	Valid	165	100.0
	Excluded(a)	0	.0
	Total	165	100.0

a Listwise deletion based on all variables in the procedure.

**Reliability Statistics**

Cronbach's Alpha	N of Items
.475	2

**Intraclass Correlation Coefficient**

	Intraclass Correlation(a)	95% Confidence Interval		F Test with True Value 0			
		Lower Bound	Upper Bound	Value	df1	df2	Sig
Single Measures	.312(b)	.167	.444	1.903	164	164	.000
Average Measures	.476	.287	.615	1.903	164	164	.000

Two-way random effects model where both people effects and measures effects are random.

a Type A intraclass correlation coefficients using an absolute agreement definition.

b The estimator is the same, whether the interaction effect is present or not.

**Reliability ICC (2,1) Consistency Definition**

Output Created	25-OCT-2006 14:23:06	
Comments		
Input	Data	E:\PhD\Systematic review\cmlsr2.sav
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	160
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics are based on all cases with valid data for all variables in the procedure.
Syntax	RELIABILITY /VARIABLES=VAR00001 VAR00002 /FORMAT=NOLABELS  /SCALE(ALPHA)=ALL/MODEL=ALPHA /ICC=MODEL(RANDOM) TYPE(CONSISTENCY) CIN=95 TESTVAL=0 .	
Resources	Elapsed Time	0:00:00.03
	Memory Available	524288 bytes
	Largest Contiguous Area	524288 bytes
	Workspace Required	128 bytes

**Case Processing Summary**

		N	%
Cases	Valid	160	100.0
	Excluded(a)	0	.0
	Total	160	100.0

a Listwise deletion based on all variables in the procedure.

**Reliability Statistics**

Cronbach's Alpha	N of Items
.487	2

**Intraclass Correlation Coefficient**

	Intraclass Correlation(a)	95% Confidence Interval		F Test with True Value 0			
		Lower Bound	Upper Bound	Value	df1	df2	Sig
Single Measures	.322(b)	.176	.454	1.948	159	159	.000
Average Measures	.487	.299	.624	1.948	159	159	.000

Two-way random effects model where both people effects and measures effects are random.

a Type C intraclass correlation coefficients using a consistency definition-the between-measure variance is excluded from the denominator variance.

b The estimator is the same, whether the interaction effect is present or not.

**APPENDIX X. TRIAL EXERCISE BOOKLET.**

**APPENDIX XI. COPIES OF ETHICS AND TRUST  
MANAGEMENT APPROVAL AND INDEMNITY LETTERS.**

**APPENDIX XII. DISC OF QUESTIONNAIRES USED IN THE  
TRIAL.**



## **APPENDIX XIII. STANDARDISED TRIAL PROCEDURES.**

This manual describes the assessment procedures for the knee trial. It is intended to supplement the assessment questionnaires and to be used as a reference manual during trial assessments. The manual is designed to enable all measurements to be taken as consistently as possible throughout the trial.

Because this is a clinical trial all the explanations to patients, instructions to patients and procedures followed need to be standardised and followed as closely as possible. Any deviations from these protocols need to be noted and explained on the patient's assessment form. Please do not give any extra verbal encouragement or give any feedback to patients regarding their performance.

Procedures are included for the following:

1. Measuring Joint Range of Motion.
2. Timed Walk Test.
3. Leg Extensor Press.
4. Timed Sit-to-Stand.

### **JOINT RANGE OF MOTION.**

Knee joint range of motion is being measured using the approach of Norkin and White (2003).

**Explanation to Participants:** The standardised explanation provided to participants is as follows:

“I am going to measure the amount of movement you have at your knee joint. The instrument I will be using to obtain the measurements is called a goniometer. It is similar to a protractor but it has two extensions called arms (*shows it to patient*). In order to obtain accurate measurements I will need to identify some anatomical landmarks (*obtain permission to remove trousers/lift skirt if necessary*). Also I'll need to press my fingers against your skin to locate some of the landmarks if that is alright? (*obtain permission*). There is a recommended testing position to help make joint measurements easier and more accurate so, if possible, I'd like you to lie down on this plinth on your back with your leg out straight. If you need any help please let me know. Also, please let me know if you are in any discomfort or wish me to stop at any time. Do you have any questions?”

**Instructions to Patient:** Each movement needs to be demonstrated as a passive movement on the participant prior to them performing the movement actively. Before each passive movement the examiner instructs the participant that “I will ask you to move your knee in exactly the same way that I move your knee”. After the passive movement then ask the patient whether they have any questions before they do the movement actively.

**Procedure:** Please measure the patient's knee joint range of motion (in degrees) using the Goniometer marked “DoH KNEE TRIAL”.

#### ***Knee Flexion.***

***Starting position:*** Participant is supine with the knee extended. The hip should be in neutral at the start of the movement. A rolled up towel is placed under the ankle to allow the knee to fully extend.

**Movement demonstration:** The examiner holds the ankle with one hand and the posterior thigh with the other. The hip is flexed to 90 degrees and stabilised prior to fully flexing the knee until end feel. The hand on the femur is also used to stabilise against hip flexion, extension or rotation.

**Goniometric alignment:** The fulcrum of the goniometer is centred over the lateral femoral epicondyle. The proximal arm is aligned with the lateral midline of the femur, using the greater trochanter as a reference point. The distal arm is aligned with the lateral midline of the fibula, using the lateral malleolus as a reference point.

**Movement:** The participant actively flexes the knee.

**Recording:** The examiner reads the goniometer and records the measurement on the data collection form in the patient assessment questionnaire.

Please remember to write in the whole range (e.g. 0- 90, or 20-120 etc) of active movement.

### ***Knee Extension.***

**Starting position:** Norkin and White (2003) state that extension is not usually measured since it is the starting position for flexion. However, it is customary clinical and research practice to record knee extension, to allow clinical conditions such as a fixed flexion deformity to be estimated, and therefore this measurement will be included in this trial. Again the starting position of supine lying will be used for the trial. The hip should aim to be in neutral position at the start of the movement. A rolled up towel is placed under the ankle to allow the knee to fully extend.

**Movement demonstration:** The examiner uses one hand to extend the knee until end feel and the other hand to stabilise the femur to prevent hip flexion, extension or rotation.

**Goniometric alignment:** The fulcrum of the goniometer is centred over the lateral femoral epicondyle. The proximal arm is aligned with the lateral midline of the femur, using the greater trochanter as a reference point. The distal arm is aligned with the lateral midline of the fibula, using the lateral malleolus as a reference point.

**Movement:** The participant actively extends the knee.

**Recording:** The examiner reads the goniometer and records the measurement on the data collection form in the patient assessment questionnaire. Please indicate whether a fixed flexion deformity is present and, if so, the amount (in degrees) Although hyperextension is uncommon in this clinical group, please indicate if this is present and, if so, the amount (in degrees).

### **TIMED WALK TEST.**

**Explanation to Patients:** The standardised explanation provided to participants is as follows:

“In a moment I’m going to ask you to walk between these two markers (*show* them) and time how long it takes you”.

**Instructions to Patients:** Patients are instructed to “Please walk at a normal comfortable speed”. I’m going to count down three two one Go! And I’d like you to start walking as soon as I say go! Is that alright? (*gain permission*)Do you have any questions? Are you ready?” Please do not speak to – or encourage – the patient during the test itself.

### **Procedure:**

Please take the patient to the marked 10m walkway. The walkway has the start and end of the 10 m length clearly marked. Please ensure the patient has their shoes on for the duration of this test. Please do not assist the patient during the test (unless a safety issue arises – in which case the test will be voided and repeated if possible).

Please check whether or not the patient usually uses a walking aid. If so, please ensure they use it for the test and please ensure that you have completed which type of walking aid they use in the patient assessment questionnaire.

Please make sure you are behind the patient during the test (to avoid pacing).

Please use a count down of “three, two, one, go!” and start timing on “go!” Stop timing when the patient crosses the 10m mark.

Please record the time taken in the patient assessment questionnaire.

### **LEG EXTENSOR PRESS.**

**Explanation to Patients:** The standardised explanation provided to participants is as follows: “Next I’m going to measure the power you have in your legs. I’m going to ask you to sit down, put one leg onto the foot plate, fold your arms and push the foot pedal down as hard and as fast as you can (*demonstrate*). There is a lot of previous research on this machine and we know that it is safe to use after knee joint replacements. We’ll practice the movement first so you can get the feel of it. And we’ll measure the power of both legs. Is that alright? (*gain permission*) Have you any questions?”

#### **Procedure:**

Sit patient on LEP seat, making sure their back is positioned against the back rest. Please ensure they are wearing shoes.

**Baseline measurement.** Adjust the seating position by placing one foot on the fully depressed footplate and pushing the seat back slowly until the leg is extended as fully as comfortable possible. Make sure the seat position is clamped in place after adjustment.

Please record the seat position in centimetres (up to 2 decimal places; eg 45.12cms) on the questionnaire. Measure the distance from the front base of the seat (labelled yellow marker 1) to the back end of the runner (labelled yellow marker 2).

**Follow Up measurements.** Each follow up measurement MUST have the same seat position. Please set the seat position to be exactly the same as the baseline measurement. Please refer to the **LEP Master List** for this measure if it is not marked on questionnaire (housed in the file marked LEP Master list in the front of the top drawer of Catherine Minns Lowe’s filing cabinet).

Please obtain measurements for both legs, starting with the leg NOT undergoing knee replacement first. Reset the recording device prior to each test.

Instruct the patient to place the foot of the leg about to be measured upon the footplate with their free foot resting on the floor.

Instruct them to “fold their arms” and ensure patients keep their arms folded throughout the tests.

Allow 2 warm up attempts for each leg. Please instruct the patient to “push the foot pedal as hard and as fast as possible after the count down of three, two, one, push!”

Allow a relaxation period of 15 seconds between each test.

Please record all the results (in Watts) in the table in the patient assessment questionnaire.

Please ensure that at least five further tests (not including the warm ups) are carried out on each leg and recorded. If the patient is still showing signs of improvement please allow them to continue up to a maximum of 10 tests per leg.

Please also make sure you have weighed the patient on the calibrated SECA bathroom scales, and recorded this measure in the patient assessment questionnaire since this is needed to perform data analysis.

### **TIMED SIT-TO-STAND.**

**Explanation to Patients:** The standardised explanation provided to participants is as follows:

“This is the last test. I’m going to ask you to sit down on this chair like this (*demonstrate*) and ask if you can stand up and sit down like this (*demonstrate*). Do you have any questions? Is that alright? (*obtain permission*)”

**Instructions to Patients:** The starting position is in sitting. Patients are instructed to “please fold your arms across your chest. In a moment I’m going to ask you to stand up and sit down as many times as you can within thirty seconds. I’m going to count down three two one Go! And I’d like you to start as soon as I say go. Do you have any questions? Are you ready?” Please do not speak to the patient – or encourage them – during the test itself.

### **Procedure:**

Please use the chair labelled “DoH KNEE TRIAL CHAIR” for use in this part of the study. Place the chair next to a wall, with the back legs braced against the wall. Please ensure the patient has their shoes on for the duration of this test. The patient starts the test in the sitting position.

Use a count down of “three, two, one, go!” Please start timing on “go!” Please count the number of times the patient completes a sit to stand and say “stop!” after thirty seconds and stop timing. Please do not assist the patient during the test (unless a safety issue arises in which case the test will be voided and repeated if appropriate).

Please record the number of completed sit-to-stands in the patient assessment form. If the patient is unable to do any sit-to-stands then please put 0 in the patient assessment questionnaire.

## APPENDIX IV. HOME VISIT CASE REPORT FORMS.

Date of visit:

Time of visit:

Duration of visit (in minutes):

Mileage (round trip):

Treatment visit no: one / two (*delete as appropriate*)

### CONTENT OF VISIT

Please fill in the approximate time spent on each of the following activities during the visit:

Component	Time (in minutes)
Subjective patient assessment	
Objective patient assessment	
Gait re-education	
Teaching new exercises	
Checking exercises	
Progressing exercises	
Giving Advice	
Task training – getting in /out of car	
Task training – getting up from chair at table	
Task training – walking outside	
Task training - stairs	
Addressing individual patient concerns	
Other: <i>please list</i>	

*Please turn over*

Please list any individual concerns/queries raised by the patient in the space below:

**APPENDIX XV. DISC OF STATISTICAL TESTS USED IN  
THE TRIAL.**

**APPENDIX XVI. COPIES OF KNEE AND HIP SYSTEMATIC  
REVIEW PUBLICATIONS.**