

Autologous chondrocyte implantation in the knee: systematic review and economic evaluation

*Hema Mistry, Martin Connock, Joshua Pink, Deepson Shyangdan,
Christine Clar, Pamela Royle, Rachel Court, Leela C Biant,
Andrew Metcalfe and Norman Waugh*



**National Institute for
Health Research**

Autologous chondrocyte implantation in the knee: systematic review and economic evaluation

Hema Mistry,¹ Martin Connock,¹ Joshua Pink,¹ Deepson Shyangdan,¹ Christine Clar,¹ Pamela Royle,¹ Rachel Court,¹ Leela C Biant,² Andrew Metcalfe³ and Norman Waugh^{1*}

¹Warwick Evidence, Division of Health Sciences, University of Warwick, Coventry, UK

²Department of Trauma and Orthopaedic Surgery, University of Manchester, Manchester, UK

³Warwick Clinical Trials Unit, University of Warwick, Coventry, UK

*Corresponding author

Declared competing interests of authors: Leela C Biant has had institutional research support from Sanofi-aventis (not related to autologous chondrocyte implantation), but has no personal conflict of interest.

Published February 2017

DOI: 10.3310/hta21060

This report should be referenced as follows:

Mistry H, Connock M, Pink J, Shyangdan D, Clar C, Royle P, *et al.* Autologous chondrocyte implantation in the knee: systematic review and economic evaluation. *Health Technol Assess* 2017;**21**(6).

Health Technology Assessment is indexed and abstracted in *Index Medicus/MEDLINE*, *Excerpta Medica/EMBASE*, *Science Citation Index Expanded (SciSearch®)* and *Current Contents®/Clinical Medicine*.

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 4.058

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, The Cochrane Library and the ISI Science Citation Index.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: nhredit@southampton.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

The HTA programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

For more information about the HTA programme please visit the website: <http://www.nets.nihr.ac.uk/programmes/hta>

This report

The research reported in this issue of the journal was commissioned and funded by the HTA programme on behalf of NICE as project number 13/65/01. The protocol was agreed in May 2014. The assessment report began editorial review in January 2015 and was accepted for publication in June 2016. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health.

© Queen's Printer and Controller of HMSO 2017. This work was produced by Mistry *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).

Health Technology Assessment Editor-in-Chief

Professor Hywel Williams Director, HTA Programme, UK and Foundation Professor and Co-Director of the Centre of Evidence-Based Dermatology, University of Nottingham, UK

NIHR Journals Library Editor-in-Chief

Professor Tom Walley Director, NIHR Evaluation, Trials and Studies and Director of the EME Programme, UK

NIHR Journals Library Editors

Professor Ken Stein Chair of HTA Editorial Board and Professor of Public Health, University of Exeter Medical School, UK

Professor Andree Le May Chair of NIHR Journals Library Editorial Group (EME, HS&DR, PGfAR, PHR journals)

Dr Martin Ashton-Key Consultant in Public Health Medicine/Consultant Advisor, NETSCC, UK

Professor Matthias Beck Chair in Public Sector Management and Subject Leader (Management Group), Queen's University Management School, Queen's University Belfast, UK

Professor Aileen Clarke Professor of Public Health and Health Services Research, Warwick Medical School, University of Warwick, UK

Dr Tessa Crilly Director, Crystal Blue Consulting Ltd, UK

Dr Eugenia Cronin Senior Scientific Advisor, Wessex Institute, UK

Ms Tara Lamont Scientific Advisor, NETSCC, UK

Professor William McGuire Professor of Child Health, Hull York Medical School, University of York, UK

Professor Geoffrey Meads Professor of Health Sciences Research, Health and Wellbeing Research Group, University of Winchester, UK

Professor John Norrie Chair in Medical Statistics, University of Edinburgh, UK

Professor John Powell Consultant Clinical Adviser, National Institute for Health and Care Excellence (NICE), UK

Professor James Raftery Professor of Health Technology Assessment, Wessex Institute, Faculty of Medicine, University of Southampton, UK

Dr Rob Riemsma Reviews Manager, Kleijnen Systematic Reviews Ltd, UK

Professor Helen Roberts Professor of Child Health Research, UCL Institute of Child Health, UK

Professor Jonathan Ross Professor of Sexual Health and HIV, University Hospital Birmingham, UK

Professor Helen Snooks Professor of Health Services Research, Institute of Life Science, College of Medicine, Swansea University, UK

Professor Jim Thornton Professor of Obstetrics and Gynaecology, Faculty of Medicine and Health Sciences, University of Nottingham, UK

Professor Martin Underwood Director, Warwick Clinical Trials Unit, Warwick Medical School, University of Warwick, UK

Please visit the website for a list of members of the NIHR Journals Library Board:
www.journalslibrary.nihr.ac.uk/about/editors

Editorial contact: nihredit@southampton.ac.uk

Abstract

Autologous chondrocyte implantation in the knee: systematic review and economic evaluation

Hema Mistry,¹ Martin Connock,¹ Joshua Pink,¹ Deepson Shyangdan,¹ Christine Clar,¹ Pamela Royle,¹ Rachel Court,¹ Leela C Biant,² Andrew Metcalfe³ and Norman Waugh^{1*}

¹Warwick Evidence, Division of Health Sciences, University of Warwick, Coventry, UK

²Department of Trauma and Orthopaedic Surgery, University of Manchester, Manchester, UK

³Warwick Clinical Trials Unit, University of Warwick, Coventry, UK

*Corresponding author norman.waugh@warwick.ac.uk

Background: The surfaces of the bones in the knee are covered with articular cartilage, a rubber-like substance that is very smooth, allowing frictionless movement in the joint and acting as a shock absorber. The cells that form the cartilage are called chondrocytes. Natural cartilage is called hyaline cartilage. Articular cartilage has very little capacity for self-repair, so damage may be permanent. Various methods have been used to try to repair cartilage. Autologous chondrocyte implantation (ACI) involves laboratory culture of cartilage-producing cells from the knee and then implanting them into the chondral defect.

Objective: To assess the clinical effectiveness and cost-effectiveness of ACI in chondral defects in the knee, compared with microfracture (MF).

Data sources: A broad search was done in MEDLINE, EMBASE, The Cochrane Library, NHS Economic Evaluation Database and Web of Science, for studies published since the last Health Technology Assessment review.

Review methods: Systematic review of recent reviews, trials, long-term observational studies and economic evaluations of the use of ACI and MF for repairing symptomatic articular cartilage defects of the knee. A new economic model was constructed. Submissions from two manufacturers and the ACTIVE (Autologous Chondrocyte Transplantation/Implantation Versus Existing Treatment) trial group were reviewed. Survival analysis was based on long-term observational studies.

Results: Four randomised controlled trials (RCTs) published since the last appraisal provided evidence on the efficacy of ACI. The SUMMIT (Superiority of Matrix-induced autologous chondrocyte implant versus Microfracture for Treatment of symptomatic articular cartilage defects) trial compared matrix-applied chondrocyte implantation (MACI[®]) against MF. The TIG/ACT/01/2000 (TIG/ACT) trial compared ACI with characterised chondrocytes against MF. The ACTIVE trial compared several forms of ACI against standard treatments, mainly MF. In the SUMMIT trial, improvements in knee injury and osteoarthritis outcome scores (KOOSs), and the proportion of responders, were greater in the MACI group than in the MF group. In the TIG/ACT trial there was improvement in the KOOS at 60 months, but no difference between ACI and MF overall. Patients with onset of symptoms < 3 years' duration did better with ACI. Results from ACTIVE have not yet been published. Survival analysis suggests that long-term results are better with ACI than with MF. Economic modelling suggested that ACI was cost-effective compared with MF across a range of scenarios.

Limitations: The main limitation is the lack of RCT data beyond 5 years of follow-up. A second is that the techniques of ACI are evolving, so long-term data come from trials using forms of ACI that are now superseded. In the modelling, we therefore assumed that durability of cartilage repair as seen in studies of older forms of ACI could be applied in modelling of newer forms. A third is that the high list prices of chondrocytes are reduced by confidential discounting. The main research needs are for longer-term follow-up and for trials of the next generation of ACI.

Conclusions: The evidence base for ACI has improved since the last appraisal by the National Institute for Health and Care Excellence. In most analyses, the incremental cost-effectiveness ratios for ACI compared with MF appear to be within a range usually considered acceptable. Research is needed into long-term results of new forms of ACI.

Study registration: This study is registered as PROSPERO CRD42014013083.

Funding: The National Institute for Health Research Health Technology Assessment programme.

Contents

| | |
|---|--------------|
| List of tables | xi |
| List of figures | xv |
| List of boxes | xix |
| List of abbreviations | xxi |
| Plain English summary | xxiii |
| Scientific summary | xxv |
| Chapter 1 Background | 1 |
| History | 1 |
| Chondral injuries | 2 |
| Interventions | 2 |
| <i>Lavage and debridement</i> | 2 |
| <i>Autologous chondrocyte implantation</i> | 3 |
| <i>Microfracture</i> | 4 |
| <i>Mosaicplasty</i> | 5 |
| <i>Conservative management</i> | 5 |
| Decision problem | 6 |
| <i>Patient group</i> | 7 |
| Chapter 2 Clinical effectiveness | 9 |
| Systematic reviews | 9 |
| <i>Inclusion criteria</i> | 9 |
| <i>Results</i> | 10 |
| Trials | 17 |
| <i>Methods</i> | 17 |
| <i>Basad et al. 2010</i> | 20 |
| <i>Saris et al. 2014</i> | 22 |
| Summary of European Medicines Agency European Public Assessment Report | 25 |
| Chapter 3 Survival analysis | 27 |
| Background | 27 |
| Methods | 28 |
| <i>Search strategy</i> | 28 |
| Results | 28 |
| Results of autologous chondrocyte implantation | 31 |
| Results of microfracture | 32 |
| Time-to-failure studies | 33 |
| <i>Caveats</i> | 33 |
| <i>Methods</i> | 33 |
| <i>Description of time-to-failure data</i> | 34 |
| <i>Results of autologous chondrocyte implantation trials</i> | 36 |
| <i>Results of autologous chondrocyte implantation observational studies</i> | 36 |
| <i>Nawaz et al. 2014 study of UK patients</i> | 40 |

| | |
|---|-----------|
| <i>Minas et al. 2014 study</i> | 45 |
| <i>Studies of failure after microfracture (Layton 2015, Knutsen 2007, Saris 2009)</i> | 45 |
| <i>Comparison of failure after autologous chondrocyte implantation and microfracture</i> | 47 |
| <i>Summary of longer-term time-to-event evidence of treatment failure after autologous chondrocyte implantation and microfracture</i> | 49 |
| <i>Pooling time-to-failure studies</i> | 49 |
| Chapter 4 Systematic review of existing economic studies for autologous chondrocyte implantation | 53 |
| Introduction | 53 |
| Methods | 53 |
| Results | 53 |
| <i>Study design, intervention and patient populations</i> | 55 |
| <i>Time horizon and length of follow-up</i> | 55 |
| <i>Study perspective and outcomes</i> | 58 |
| <i>Resource use, costs and discounting</i> | 58 |
| <i>Base-case results and sensitivity analyses</i> | 59 |
| <i>Limitations of the studies</i> | 59 |
| <i>Quality assessment</i> | 60 |
| Discussion | 60 |
| Chapter 5 Commentary on submissions by manufacturers and by the Oswestry group, including data from the ACTIVE trial | 63 |
| ChondroSelect | 63 |
| <i>Trial data: autologous chondrocyte implantation–periosteal flap versus microfracture: TIG/ACT/01/2000</i> | 63 |
| <i>Case series</i> | 66 |
| <i>Registry cohort</i> | 67 |
| <i>Belgian reimbursement scheme</i> | 67 |
| <i>Cost-effectiveness</i> | 67 |
| Aastrom Biosciences submission | 71 |
| <i>Cost-effectiveness</i> | 71 |
| Submission by OsCell | 72 |
| <i>The ACTIVE trial</i> | 72 |
| <i>Cost-effectiveness</i> | 73 |
| Discussion: clinical effectiveness | 75 |
| <i>Previous repairs</i> | 75 |
| <i>Duration of symptoms</i> | 76 |
| <i>ACI-C or MACI?</i> | 76 |
| <i>Predicting success</i> | 76 |
| <i>Survival of repairs</i> | 76 |
| Commentary on SoBi submission October 2015: survival analysis | 77 |
| Chapter 6 The cost-effectiveness of autologous chondrocyte implantation | 83 |
| Introduction | 83 |
| Model structure | 83 |
| <i>Clinical pathways</i> | 84 |
| <i>Knee repairs</i> | 85 |
| <i>Knee replacements</i> | 86 |
| <i>Deaths</i> | 86 |
| <i>Markov model structure</i> | 87 |
| <i>Base-case analysis</i> | 87 |

| | |
|---|------------|
| Model inputs | 87 |
| <i>Transition probabilities</i> | 87 |
| <i>Utilities</i> | 89 |
| <i>Resource use and costs</i> | 91 |
| <i>Complications</i> | 93 |
| <i>Mortality</i> | 93 |
| <i>Measuring cost-effectiveness</i> | 93 |
| Scenario and sensitivity analyses | 94 |
| Results | 94 |
| <i>Base-case cost-effectiveness results</i> | 94 |
| <i>Scenario and sensitivity analysis cost-effectiveness results</i> | 97 |
| Discussion | 118 |
| <i>Methods and summary of findings</i> | 119 |
| <i>Strengths and limitations</i> | 120 |
| Additional analyses based on survival analysis | 121 |
| <i>New base case</i> | 121 |
| <i>Sensitivity analyses on prices</i> | 122 |
| <i>Sensitivity analyses: post-repair utility</i> | 124 |
| <i>Subgroup analyses</i> | 128 |
| <i>Pooled autologous chondrocyte implantation curve (six studies)</i> | 132 |
| Using utility data from Vericel | 134 |
| Summary of long-study modelling | 135 |
| Chapter 7 Discussion | 137 |
| Strengths and weaknesses of evidence | 137 |
| Other issues | 138 |
| <i>Asymptomatic lesions</i> | 138 |
| <i>Osteoarthritis</i> | 138 |
| <i>Age threshold for knee replacement</i> | 140 |
| <i>Body mass index</i> | 140 |
| <i>Research needs</i> | 140 |
| <i>Fourth-generation autologous chondrocyte implantation</i> | 141 |
| <i>New forms of microfracture</i> | 142 |
| <i>Metal or plastic patches for knees</i> | 142 |
| <i>Conclusion</i> | 142 |
| Acknowledgements | 143 |
| References | 145 |
| Appendix 1 Search strategies for systematic review and primary studies | 161 |
| Appendix 2 Flow diagram systematic review | 165 |
| Appendix 3 Quality assessment of reviews | 167 |
| Appendix 4 Studies included in reviews | 169 |
| Appendix 5 Characteristics of systematic reviews | 173 |
| Appendix 6 Results and conclusions of systematic reviews | 187 |
| Appendix 7 Excluded studies: survival analysis | 201 |

| | |
|--|------------|
| Appendix 8 Included long-term studies: data extraction and quality assessment | 203 |
| Appendix 9 Models of time to failure included published autologous chondrocyte implantation and microfracture studies | 251 |
| Appendix 10 Economic search strategies | 285 |
| Appendix 11 Annual transition probabilities | 287 |

List of tables

| | | |
|-----------------|--|-----------|
| TABLE 1 | The evolution of ACI | 4 |
| TABLE 2 | Outcomes used in cartilage repair studies | 19 |
| TABLE 3 | Reason for exclusion of studies | 20 |
| TABLE 4 | Response rates after prior cartilage procedures | 24 |
| TABLE 5 | Included studies | 29 |
| TABLE 6 | Time-to-event data presented in relevant studies | 35 |
| TABLE 7 | Failure definitions used in relevant studies | 35 |
| TABLE 8 | Baseline characteristics reported for patients in six ACI studies | 38 |
| TABLE 9 | Patient characteristics in the MF arms of the Saris and Knutsen RCTs | 46 |
| TABLE 10 | Failure of ACI and MF according to previous intervention (data from Vanlauwe <i>et al.</i> 2011) | 48 |
| TABLE 11 | Information criteria for models to pooled MF data from Knutsen and Vanlauwe | 51 |
| TABLE 12 | Characteristics of previous economic studies | 56 |
| TABLE 13 | Mean change in overall KOOS and subscales from baseline at 60 months | 65 |
| TABLE 14 | Mean change in overall KOOS from baseline | 65 |
| TABLE 15 | Comparison of baseline characteristics trial and case series patients | 66 |
| TABLE 16 | Patients' responses | 73 |
| TABLE 17 | Costs of ACI and its comparators by Oswestry | 74 |
| TABLE 18 | The EQ-5D scores in the ACTIVE trial | 75 |
| TABLE 19 | Increases in EQ-5D from baseline | 75 |
| TABLE 20 | Information criteria for six studies with ACI arms examined in the Assessment Group report | 77 |
| TABLE 21 | Base-case mean utility values used in the economic model | 90 |
| TABLE 22 | Base-case mean costs (UK pounds sterling) used in the economic model | 92 |
| TABLE 23 | Base-case resource use for economic model | 92 |

| | |
|--|------------|
| TABLE 24 Base-case deterministic and probabilistic cost-effectiveness results (by scenario) | 95 |
| TABLE 25 Base-case deterministic and probabilistic cost-effectiveness results | 96 |
| TABLE 26 Sensitivity cost-effectiveness results (by scenario): cell cost reduction | 97 |
| TABLE 27 Sensitivity cost-effectiveness results: cell cost reduction | 101 |
| TABLE 28 Sensitivity cost-effectiveness results (by scenario): changing the time horizon | 103 |
| TABLE 29 Sensitivity cost-effectiveness results: changing the time horizon | 108 |
| TABLE 30 Sensitivity cost-effectiveness results (by scenario): MF procedure as a day case surgery and not as an inpatient | 111 |
| TABLE 31 Sensitivity cost-effectiveness results: MF procedure as a day case surgery and not as an inpatient | 113 |
| TABLE 32 Sensitivity cost-effectiveness results (by scenario): improving the success rates of MF | 114 |
| TABLE 33 Sensitivity cost-effectiveness results: improving the success rates of MF | 116 |
| TABLE 34 Sensitivity cost-effectiveness results (by scenario): starting age for cohort is 45 years | 117 |
| TABLE 35 Sensitivity cost-effectiveness results: starting age for cohort is 45 years | 119 |
| TABLE 36 Base-case deterministic and probabilistic cost-effectiveness results | 122 |
| TABLE 37 Deterministic and probabilistic cost-effectiveness results: £6000 price | 123 |
| TABLE 38 Deterministic and probabilistic cost-effectiveness results: £8000 price | 124 |
| TABLE 39 Deterministic and probabilistic cost-effectiveness results: £12,000 price | 125 |
| TABLE 40 Deterministic and probabilistic cost-effectiveness results: utility = 0.654 | 126 |
| TABLE 41 Deterministic and probabilistic cost-effectiveness results: utility = 0.817 | 127 |
| TABLE 42 Deterministic and probabilistic cost-effectiveness results: utility = 0.746 | 128 |
| TABLE 43 Deterministic and probabilistic cost-effectiveness results: previous intervention | 129 |
| TABLE 44 Deterministic and probabilistic cost-effectiveness results: no previous intervention | 130 |
| TABLE 45 Deterministic cost-effectiveness results: Kellgren–Lawrence grade 0 | 131 |
| TABLE 46 Deterministic cost-effectiveness results: Kellgren–Lawrence grade 1 | 132 |

| | | |
|-----------------|---|------------|
| TABLE 47 | Deterministic cost-effectiveness results: Kellgren–Lawrence grade 2 | 132 |
| TABLE 48 | Deterministic cost-effectiveness results: Kellgren–Lawrence grade 3 | 133 |
| TABLE 49 | Deterministic and probabilistic cost-effectiveness results: six ACI data sets | 133 |
| TABLE 50 | Utility data from Vericel | 134 |
| TABLE 51 | Deterministic and probabilistic cost-effectiveness results | 134 |
| TABLE 52 | Primary prospective comparative studies in reviews and from extra searches | 170 |
| TABLE 53 | Characteristics and quality of systematic reviews | 173 |
| TABLE 54 | Results and conclusions of systematic reviews | 187 |
| TABLE 55 | Systematic review conclusions | 197 |
| TABLE 56 | The AIC and BIC information criteria for parametric models | 251 |
| TABLE 57 | Annual transition probabilities: 20–54 years | 291 |
| TABLE 58 | Annual transition probabilities: 55+ years | 292 |
| TABLE 59 | Annual transition probabilities: 55+ years (for all scenarios) | 293 |
| TABLE 60 | Sensitivity analysis: mean utility values used in the economic model | 293 |

List of figures

| | |
|--|-----------|
| FIGURE 1 Reconstructed KM plots (95% CI) of time to failure in two RCTs | 36 |
| FIGURE 2 Extrapolated Weibull distributions fit to reconstructed IPD from the RCT studies | 37 |
| FIGURE 3 (a) Reconstructed KM plots for ACI studies; and (b) best parametric fit (95% CIs) and gamma models for six ACI arms | 39 |
| FIGURE 4 Reconstructed KM plot (a) and extrapolated log-logistic model; and (b) for the Nawaz <i>et al.</i> whole cohort | 41 |
| FIGURE 5 Reconstructed plots for Nawaz <i>et al.</i> Previous and no previous intervention subgroups: (a) log-normal models; and (b) Weibull models | 42 |
| FIGURE 6 Reconstructed KM plots and log-normal models for time to failure according to site of intervention, Nawaz <i>et al.</i> (a) KM according to intervention site; and (b) subgroups by intervention site log-normal models based on HRs | 43 |
| FIGURE 7 Time to failure by grade of OA | 44 |
| FIGURE 8 Effect of previous interventions on failure rates | 45 |
| FIGURE 9 Best fit models to observed MF failure in three studies | 47 |
| FIGURE 10 Modelled failure profiles following ACI or MF | 47 |
| FIGURE 11 Modelled failure in MF studies compared with worst-performing subgroups from Nawaz <i>et al.</i> ACI study | 48 |
| FIGURE 12 Time to failure after pooling four ACI studies and three MF studies | 50 |
| FIGURE 13 Time to failure after pooling four ACI and two MF studies (Knutsen and Saris) | 50 |
| FIGURE 14 Time to failure after pooling four ACI and two MF studies (Knutsen and Vanlauwe) | 51 |
| FIGURE 15 A PRISMA flow diagram for cost-effectiveness studies | 54 |
| FIGURE 16 Pooled ACI studies compared with SoBi curves | 78 |
| FIGURE 17 Alternative models for six pooled ACI studies ($n = 1270$) | 79 |
| FIGURE 18 SoBi hybrid curve for ACI failure | 79 |
| FIGURE 19 Exponential and other parametric models of MF failure based on the MF arm of the TIG/ACT study | 80 |

| | |
|---|------------|
| FIGURE 20 Patient pathways for ACI or MF: scenario 1 (above), scenario 2 (below) | 84 |
| FIGURE 21 Clinical pathways for patients with articular cartilage defects of the knee joint | 85 |
| FIGURE 22 Markov model structure for patients with articular cartilage defects of the knee joint | 88 |
| FIGURE 23 Proportion of patients achieving success/failure with ACI or MF at 36 months | 89 |
| FIGURE 24 Cost-effectiveness acceptability curve: base case results: (a) scenario 1; and (b) scenario 2 | 96 |
| FIGURE 25 Cost-effectiveness acceptability curve: base-case results – all sequences | 97 |
| FIGURE 26 Cost-effectiveness acceptability curves: cost reductions: (a) 25% cell cost reduction – scenario 1; (b) 25% cell cost reduction – scenario 2; (c) 50% cell cost reduction – scenario 1; (d) 50% cell cost reduction – scenario 2; (e) 75% cell cost reduction – scenario 1; and (f) 75% cell cost reduction – scenario 2 | 99 |
| FIGURE 27 Cost-effectiveness acceptability curves: cell cost reduction: (a) 25% cell cost reduction; (b) 50% cell cost reduction; and (c) 75% cell cost reduction | 102 |
| FIGURE 28 Cost-effectiveness acceptability curves: different time horizons: (a) 10-year time horizon – scenario 1; (b) 10-year time horizon – scenario 2; (c) 20-year time horizon – scenario 1; (d) 20-year time horizon – scenario 2; (e) 30-year time horizon – scenario 1; (f) 30-year time horizon – scenario 2; (g) 40-year time horizon – scenario 1; (h) 40-year time horizon – scenario 2; (i) 50-year time horizon – scenario 1; and (j) 50-year time horizon – scenario 2 | 105 |
| FIGURE 29 Cost-effectiveness acceptability curve: changing the time horizon | 110 |
| FIGURE 30 Cost-effectiveness acceptability curves: MF as a day case | 112 |
| FIGURE 31 Cost-effectiveness acceptability curve: MF as a day case procedure | 113 |
| FIGURE 32 Cost-effectiveness acceptability curves: increase in MF success rate: (a) 20% increase in MF success rate – scenario 1; (b) 20% increase in MF success rate – scenario 2; (c) 40% increase in MF success rate – scenario 1; and (d) 40% increase in MF success rate – scenario 2 | 115 |
| FIGURE 33 Cost-effectiveness acceptability curves: increase in MF success rate | 117 |
| FIGURE 34 Cost-effectiveness acceptability curves: starting age is 45 years: (a) scenario 1; and (b) scenario 2 | 118 |
| FIGURE 35 Cost-effectiveness acceptability curve: starting age of cohort is 45 years | 119 |
| FIGURE 36 Cost-effectiveness acceptability curve: base case | 122 |
| FIGURE 37 Cost-effectiveness acceptability curve: £6000 | 123 |

| | |
|---|------------|
| FIGURE 38 Cost-effectiveness acceptability curve: £8000 | 124 |
| FIGURE 39 Cost-effectiveness acceptability curve: £12,000 | 125 |
| FIGURE 40 Cost-effectiveness acceptability curve: utility = 0.654 | 126 |
| FIGURE 41 Cost-effectiveness acceptability curve: utility = 0.817 | 127 |
| FIGURE 42 Cost-effectiveness acceptability curve: utility = 0.746 | 128 |
| FIGURE 43 Cost-effectiveness acceptability curve (previous interventions) | 129 |
| FIGURE 44 Cost-effectiveness acceptability curve: no previous interventions | 130 |
| FIGURE 45 Cost-effectiveness acceptability curve: six ACI data sets | 133 |
| FIGURE 46 Cost-effectiveness acceptability curve | 135 |
| FIGURE 47 Flow diagram for long-term studies | 164 |
| FIGURE 48 A PRISMA study flow diagram for searches for systematic reviews | 165 |
| FIGURE 49 Bentley <i>et al.</i> 2012 ACI arm model fits | 258 |
| FIGURE 50 Biant <i>et al.</i> 2014 ACI model fits | 259 |
| FIGURE 51 Knutsen <i>et al.</i> 2007 model fits, ACI arm | 260 |
| FIGURE 52 Knutsen <i>et al.</i> 2007 model fits, MF arm | 261 |
| FIGURE 53 Layton <i>et al.</i> 2015 model fits, MF | 262 |
| FIGURE 54 Minas <i>et al.</i> 2014 model fits, ACI | 263 |
| FIGURE 55 Minas <i>et al.</i> 2014 model fits, ACI, no previous intervention | 264 |
| FIGURE 56 Minas <i>et al.</i> 2014 model fits, ACI, previous intervention | 265 |
| FIGURE 57 Moseley <i>et al.</i> 2010 model fits, ACI | 266 |
| FIGURE 58 Nawaz <i>et al.</i> 2014 model fits, ACI, whole cohort | 267 |
| FIGURE 59 Nawaz <i>et al.</i> 2014 model fits, ACI, previous intervention | 268 |
| FIGURE 60 Nawaz <i>et al.</i> 2014 model fits, ACI, no previous intervention | 269 |
| FIGURE 61 Nawaz <i>et al.</i> 2014 model fits, ACI, lateral femoral site | 270 |
| FIGURE 62 Nawaz <i>et al.</i> 2014 model fits, ACI, medial femoral site | 271 |
| FIGURE 63 Nawaz <i>et al.</i> 2014 model fits, ACI, multisite | 272 |
| FIGURE 64 Nawaz <i>et al.</i> 2014 model fits, ACI, patella site | 273 |

| | |
|--|------------|
| FIGURE 65 Nawaz <i>et al.</i> 2014 model fits, ACI, trochlea site | 274 |
| FIGURE 66 Nawaz <i>et al.</i> 2014 model fits, ACI, grade 0 degradative change | 275 |
| FIGURE 67 Nawaz <i>et al.</i> 2014 model fits, ACI, grade 1 degradative change | 276 |
| FIGURE 68 Nawaz <i>et al.</i> 2014 model fits, ACI, grade 2 degradative change | 277 |
| FIGURE 69 Nawaz <i>et al.</i> 2014 model fits, ACI, grade 3 degradative change | 278 |
| FIGURE 70 Niemeyer <i>et al.</i> 2014 model fits, ACI | 279 |
| FIGURE 71 Saris <i>et al.</i> 2009, ACI arm | 280 |
| FIGURE 72 Saris <i>et al.</i> 2009, MF arm | 281 |
| FIGURE 73 Vanlauwe <i>et al.</i> 2011, MF arm model fits; IPD reconstructed without risk table data | 282 |
| FIGURE 74 Vanlauwe <i>et al.</i> 2011, ACI arm model fits | 283 |

List of boxes

| | |
|---|-----------|
| BOX 1 Clinical features of ACI-P | 3 |
| BOX 2 National Institute for Health and Care Excellence Guidance. Reproduced with permission from NICE | 5 |
| BOX 3 Licences for ChondroCelect and Verigen MACI® | 6 |
| BOX 4 Consensus statements from BASK | 7 |
| BOX 5 Autologous chondrocyte implantation: primary comparative studies in reviews | 13 |

List of abbreviations

| | | | |
|----------|---|-------|---|
| ACI | autologous chondrocyte implantation | FDA | Food and Drug Administration |
| ACI-C | autologous chondrocyte implantation–collagen cap | GP | general practitioner |
| ACI-P | autologous chondrocyte implantation–periosteal flap | HCHS | Hospital and Community Health Services |
| ACTIVE | Autologous Chondrocyte Transplantation/Implantation Versus Existing Treatment | HR | hazard ratio |
| ADL | activities of daily living | HTA | Health Technology Assessment |
| AE | adverse event | ICER | incremental cost-effectiveness ratio |
| BASK | British Association for Surgery of the Knee | ICRS | International Cartilage Repair Society |
| BMI | body mass index | IKDC | International Knee Documentation Committee |
| BSC | best supportive care | IPD | individual patient data |
| CAIS | cartilage autograft implantation system | KM | Kaplan–Meier |
| CCI | characterised chondrocyte implantation | KOOS | knee injury and osteoarthritis outcome score |
| CCT | controlled clinical trial | MACI | Matrix ACI: matrix-applied chondrocyte implantation |
| CEAC | cost-effectiveness acceptability curve | MF | microfracture |
| CGI-E | Clinical Global Impression measures of efficacy | MFF | market forces factor |
| CGI-I | Clinical Global impression measures of improvement | MRI | magnetic resonance imaging |
| CHEERS | Consolidated Health Economic Evaluation Reporting Standards | MSAC | Medical Services Advisory Committee |
| CI | confidence interval | NICE | National Institute for Health and Care Excellence |
| CPV | continuous passive motion | NIH | National Institutes of Health |
| CRD | Centre for Reviews and Dissemination | NRIG | no re-intervention group |
| DVT | deep-vein thrombosis | NSAID | non-steroidal anti-inflammatory drug |
| EMA | European Medicines Agency | OA | osteoarthritis |
| EQ-5D | EuroQol-5 Dimensions | OAT | osteocondral autograft transfer |
| EQ-5D-3L | EuroQol-5 Dimensions, three-level version | OATS | osteocondral autograft transfer system |
| | | OCD | osteocondritis dissecans |
| | | ONS | Office for National Statistics |

LIST OF ABBREVIATIONS

| | | | |
|--------|--|---------|--|
| PbR | Payment by Results | SAE | serious adverse event |
| PKR | partial knee replacement | SD | standard deviation |
| PRISMA | Preferred Reporting Items for Systematic Reviews and Meta-Analyses | SE | standard error |
| PSA | probabilistic sensitivity analysis | SF-36 | Short Form questionnaire-36 items |
| PSS | Personal Social Services | SUMMIT | Superiority of Matrix-induced autologous chondrocyte implant versus Microfracture for Treatment of symptomatic articular cartilage defects |
| QALY | quality-adjusted life-year | TA | technology appraisal |
| QoL | quality of life | TIG/ACT | TIG/ACT/01/2000 |
| RCT | randomised controlled trial | TKR | total knee replacement |
| RIG | re-intervention group | TTF | time to treatment failure |
| RJAH | Robert Jones and Agnes Hunt Hospital, Oswestry | VAS | visual analogue scale |
| RR | relative risk | WORMS | whole-organ MRI score |
| SA | sensitivity analysis | | |

Plain English summary

Damage to the cartilage that covers the ends of the knee bones is common, especially in sportspeople. The ability of cartilage to repair itself is very limited. The damage may lead to later osteoarthritis and a need for knee replacements.

Attempts to repair the damaged cartilage have mostly used a procedure called microfracture (MF), involving drilling small holes into the bone under the damaged area. A clot of blood then covers the damaged area and over time forms scar tissue called fibrocartilage. This relieves symptoms, but fibrocartilage is not as tough as natural cartilage, and wears out after a few years.

In the procedure called autologous chondrocyte implantation (ACI), a small piece of healthy cartilage is removed from the knee, and the cartilage-producing cells (chondrocytes) are cultured in the laboratory until there are many millions of them. They are then used to patch the area of damaged cartilage. The hope is that the chondrocytes will form natural cartilage, which will last longer than fibrocartilage.

We reviewed trials and other studies that reported that ACI gave better long-term results than MF. The main limitation was the lack of long-term follow-up data from modern methods of ACI. However, we thought that data on the cartilage formed using older methods of ACI could be applied to modern methods.

Our analysis suggested that ACI is cost-effective, taking into account short-term improvements in symptoms and reduced need for further repairs and, in the long term, knee replacements.

Scientific summary

Background

The surfaces of the bones in the knee are covered with articular cartilage, a rubber-like substance that is very smooth, allowing frictionless movement in the joint, and acting as a shock absorber. The cells that form the cartilage are called chondrocytes. Natural cartilage is called hyaline cartilage.

Various methods have been used to try to repair cartilage defects, usually aiming to replace the damaged cartilage using bone marrow cells, including stem cells, which then form a tissue called fibrocartilage. The commonest way of doing this is called microfracture (MF). Small holes are drilled through the bone underlying the damaged area to allow the marrow cells to fill the defect. However, the fibrocartilage formed is less durable than natural hyaline cartilage.

Autologous chondrocyte implantation (ACI) aims to replace the damaged cartilage with hyaline cartilage. A small piece of articular cartilage is taken from the knee, and the chondrocytes are cultured in the laboratory, until there are millions of cells, which are implanted into the damaged area.

The methods of ACI have evolved. In the first generation of ACI (ACI-P – 'P' for periosteum), the cultured cells were implanted as a liquid suspension, and covered with a cap made from periosteum – the tough fibrous tissue that covers bones. This required a procedure to harvest the periosteum, which caused discomfort to the patient afterwards.

In second-generation ACI, the periosteal cover was replaced by a collagen cover (ACI-C for short), but the cells were still in liquid suspension, and the cover still had to be stitched in place.

One development in ACI has been 'characterisation', a process in which the cells thought to have the best ability to form hyaline cartilage are selected during culture.

In the third generation of ACI, the cells are seeded or loaded into a collagen membrane, which is implanted into the defect. This is referred to as matrix-applied chondrocyte implantation (MACI).

Decision problem

The scope from the National Institute for Health and Care Excellence (NICE) for this appraisal mentions three forms of ACI:

1. The ChondroSelect system from TiGenix (Leuven, Belgium), in which characterised cells are capped with biodegradable collagen: ACI-C.
2. The Matrix ACI system [matrix-applied characterised autologous cultured chondrocyte implant (MACI®)] now marketed by Vericel (Cambridge, MA, USA).
3. ACI wherein the cells are cultured in hospital or research laboratories, such as the Robert Jones and Agnes Hunt Hospital in Oswestry, termed 'traditional ACI' in the NICE scope.

The main comparator is MF.

Clinical effectiveness

We reviewed previous systematic reviews of the comparative effectiveness of various forms of ACI and MF. We then searched for trials [randomised controlled trials (RCTs)] that used the most recent forms of ACI.

The reviews were mostly inconclusive on the choice between ACI and MF, for reasons including poor quality of primary studies, the heterogeneities of patients recruited, ACI methods used, outcome measures, variations in previous surgery and short follow-up periods.

Four RCTs have been published since the last appraisal provided evidence on the efficacy of ACI. These are:

Basad E, Ishaque B, Bachmann G, Stürz H, Steinmeyer J. Matrix-induced autologous chondrocyte implantation versus microfracture in the treatment of cartilage defects of the knee: a 2-year randomised study. *Knee Surg Sports Traumatol Arthrosc* 2010;**18**:519–27. <http://dx.doi.org/10.1007/s00167-009-1028-1>.

SUMMIT (Superiority of Matrix-induced autologous chondrocyte implant versus Microfracture for Treatment of symptomatic articular cartilage defects) Saris D, Price A, Widuchowski W, Bertrand-Marchand M, Caron J, Drogset JO, *et al.* Matrix-applied characterized autologous cultured chondrocytes versus microfracture: two-year follow-up of a prospective randomized trial. *Am J Sports Med* 2014;**42**:1384–94. <http://dx.doi.org/10.1177/0363546514528093>.

The TIG/ACT/01/2000 trial (hereafter TIG/ACT trial) Vanlauwe J, Saris DB, Victor J, Almqvist KF, Bellemans J, Luyten FP, TIG/ACT/01/2000&EXT Study Group. Five-year outcome of characterized chondrocyte implantation versus microfracture for symptomatic cartilage defects of the knee: early treatment matters. *Am J Sports Med* 2011;**39**:2566–74. <http://dx.doi.org/10.1177/0363546511422220>.

The ACTIVE (Autologous Chondrocyte Transplantation/Implantation Versus Existing Treatment) trial (390 patients) [Keele University. *ACTIVE Trial Web Site*. 2011. URL: www.active-trial.org.uk/ (accessed 25 July 2016)].

Two of the trials, Basad *et al.* with 60 patients and SUMMIT by Saris *et al.* with 144 patients, compared MACI with MF. The TIG/ACT trial with 118 patients compared ACI-P with characterised chondrocytes against MF. The ACTIVE trial compared several forms of ACI against standard treatment, mainly MF.

The primary outcome measures in the Basad trial were Tegner and Lysholm scores. Lysholm scores improved in both MACI and MF groups from baseline to 12 months, but the improvement was maintained to 24 months only in the MACI group (92 vs. 69; $p = 0.005$). Tegner scores improved in both groups, but more so in the MACI group.

In the SUMMIT trial, improvements in knee injury and osteoarthritis outcome knee injury and osteoarthritis outcome scores (KOOSs) were significantly greater in the MACI group than in the MF group. The proportion of responders was higher with MACI. Factors that predicted better results with MACI were male gender, a median age of < 34.5 years, presence of a single lesion due to acute trauma, history of only one previous surgical procedure, and lesion of size > 4 cm² located on the femoral condyle. More patients in the MF group reported adverse events (AEs), most frequently arthralgia.

In the TIG/ACT trial of ChondroCelect, the overall KOOS improved at 60 months with both treatments, with no statistically significant difference. Patients with onset of symptoms < 3 years' duration did better with ACI-P. More patients in the ACI-P group experienced AEs but they were mild to moderate in intensity.

The ACTIVE trial is comparing ACI (including ACI-P, ACI-C and MACI) against standard treatments (MF, abrasion, drilling, mosaicplasty).

Autologous chondrocyte implantation is less successful in patients who have had previous MF than if it is done as first repair, because MF damages the subchondral bone.

Longer-term results: survival analysis

The trials included under the original scope from NICE on second- and third-generation ACI provided results only up to 3 and 5 years, and, for modelling of cost-effectiveness, longer-term outcome data were desirable. It was decided that longer-term data from ACI-P could be used, based on an assumption that data on longer-term outcomes of chondral defect repairs from studies of ACI-P could be extrapolated to survival of repairs after ACI-C and MACI. ACI-P has been superseded because the new techniques are simpler and quicker, and because the use of periosteum required harvesting and ensuring a watertight cap, and could lead to overgrowth hypertrophy requiring reoperation and shaving of the graft, and the extra discomfort to patients from these procedures. The collagen cap is much easier to use. The third generation of ACI in which the cells are seeded onto the collagen membrane is quicker still.

It was felt that results after ACI-C and MACI would at least be no worse than after ACI-P.

We searched for studies reporting longer-term results of ACI and MF.

Survival analysis – time to failure in longer-term studies: we included six studies of long-term results of ACI, the best of which was by Nawaz *et al.* from Stanmore. (Nawaz SZ, Bentley G, Briggs TW, Carrington RW, Skinner JA, Gallagher KR, Dhinsa BS. Autologous chondrocyte implantation in the knee: mid-term to long-term results. *J Bone Joint Surg Am* 2014;**96**:824–30.) It was best because of its size (827 patients – greater than the other studies put together); because it reflected UK practice (albeit from a centre of excellence); because it provided data from the period 1998–2008, on different generations of ACI; and because it provided very useful subgroup data.

The findings of the Nawaz study include:

- ACI graft survival was 78% at 5 years and 51% at 10 years for the whole cohort.
- There was no difference between survival rates of ACI-P and ACI-C, and MACI. Most (63%) received MACI.
- Outcomes were much poorer in patients who had had previous attempts at cartilage repair such as MF, with an almost fivefold failure rate.
- The presence of osteoarthritis (OA) increased failure rates. Patients with Kellgren–Lawrence grades 2 and 3 had only 25% graft survival to 10 years.

We used the Nawaz results as the main input into survival analysis and cost-effectiveness, but also did a sensitivity analysis (SA) incorporating five other long-term studies of ACI.

There were few long-term studies of MF. We constructed survival curves based on 5-year data from only three studies: two trials with 40 and 61 patients and an observational study from routine care in the USA with 3498 patients having MF.

The ACI groups had lower failure rates than the MF cohorts, except for the ACI group with previous attempts at repair or with OA. Data were sparse on results of MF in previously treated patients.

In summary:

- More long-term evidence was available for ACI than for MF.
- Study data were generally still too short term. Only one published study allowed an estimate of observed median time to failure.
- Caveat: immaturity of failure data necessitated parametric modelling beyond observed data so as to predict lifetime failure. Such extrapolations assume that curves based on the observed data will continue.

- Most participants in most studies had had previous attempts at repair. Two ACI studies with survival analyses extending to at least 10 years reported that treatment failure was far more frequent in patients who had experienced prior intervention(s). This reduced the likelihood of success after ACI and makes extrapolation of results from older studies to ACI as first procedure rather pessimistic.
- The best fits of long-term failure after ACI were usually characterised by models that, when extrapolated beyond the observed data, indicated gradually decreasing hazard (probability of failure decreasing with time).
- Conversely, good fits to limited data available for MF were characterised by models that indicated linearly increasing hazard (probability of failure increasing with time).

Cost-effectiveness

Review of previous economic studies

We reviewed existing economic evaluations of the use of ACI and MF for repairing symptomatic articular cartilage defects of the knee. A broad search was done in MEDLINE, EMBASE, NHS Economic Evaluation Database and Web of Science for studies published since the last Health Technology Assessment review in 2005.

We found six relevant articles, all with shortcomings, most notably the lack of long-term clinical follow-up and good quality-of-life (QoL) data.

Review of submissions received

We reviewed the submissions from Swedish Orphan Biovitrum AB (SoBi, Stockholm, Sweden) on ChondroCelect, from Aastrom Biosciences (now Vericel, Cambridge, MA, USA) on MACI[®], and from OsCell, including unpublished data from the ACTIVE trial.

Swedish Orphan Biovitrum AB developed a de novo Markov economic model. Their modelling assumed that MF was the comparator; if the first repair fails then patients can have a second repair, but only with MF, and the main driver was time to failure of the first repair. They used data from the TIG/ACT trial. Their key assumptions were that fewer patients who had ACI needed second repairs and that they had a longer duration of success, thereby postponing the need for knee replacement. Their base-case incremental cost-effectiveness ratio (ICER) was about £9000 per quality-adjusted life-year (QALY).

Aastrom Biosciences did not provide any cost-effectiveness analysis but did provide a costing forecast. They explored two scenarios, one with MACI or ACI as first procedure, and the other with MF. Based on data from the SUMMIT trial, they estimated that there would be cost savings from using MACI due to the lower need for further repairs.

The Oswestry group provided a cost-effectiveness analysis for the ACTIVE trial. This analysis used EQ-5D-3L (EuroQol-5 Dimensions, three-level version) data based on up to 8 years of follow-up. It assumed a cost for cells of £4125, based on production by OsCell. The data showed little difference between ACI and MF for the first 4 years but, after that, EQ-5D results were better in the ACI group, with a cost per QALY for ACI compared with MF of around £6000.

Warwick Evidence modelling

We constructed a lifetime Markov model, starting with a cohort of people aged 33 years with symptomatic articular cartilage defects of the knee treated with ACI or MF. The analysis considered the need for subsequent events including further repairs and later knee replacements. Most patients (87.5%) did not need a second repair. We created two scenarios to allow direct comparisons: in scenario 1 all second repairs were ACI, and in scenario 2 all second repairs were MF. Secondary analyses considered other options including ACI after prior MF.

For the base-case analysis, we used data mainly from the TIG/ACT trial of ChondroCelect and the SUMMIT trial of MACI.

The results indicated that ACI is more cost-effective than MF as a first repair, and that if a second repair is needed this should also be ACI. The base-case discounted ICER for ACI compared with MF was just over £14,000 per QALY for scenario 1, and just under £16,000 per QALY for scenario 2.

Results from SAs were in line with the base-case results.

The key drivers in the base case were the cost of cells for ACI and the relative durations of benefit from ACI and MF. After the first few years (varying among studies) ACI was more beneficial (more gain in QALYs) and led to cost savings to the NHS [fewer people in need of a second repair or of a total knee replacement (TKR), and first TKR postponed reducing the need for second TKR].

Limitations in the economic analyses included uncertainties with long-term progression rates and QoL data.

We then used data from the long-term survival analysis, using the whole Nawaz cohort results for ACI, and pooling the MF results from three studies. At the request of NICE, we used an implantation cost of £2396 (assuming an inpatient stay), and we omitted the option for MF failure to be followed by another MF. So the options were:

- MF followed by ACI if another procedure was considered necessary in the short term. In the long term, patients would be considered for knee replacement, but most would still be too young for that after MF failure.
- ACI followed by MF if another attempt at repair was necessary.
- ACI followed by a second ACI if another attempt at repair was necessary.

Microfracture followed if necessary by ACI was the lowest cost option, and we compared other options with that. ACI followed by MF was dominated by ACI followed if necessary by ACI, because of the poor long-term results of MF.

The ICER for ACI as primary procedure compared with MF was around £19,000 – a little less in deterministic analysis, a bit more in probabilistic. A caveat is necessary – the marginal QALY gains were very small, at 0.0650 in deterministic and 0.0824 in probabilistic.

We carried out a range of SAs. In the base case, we used a cell cost of £16,000 in line with published prices, but we are aware of discounted prices that vary by time and place. The deterministic ICERs for ACI as first procedure compared with MF were as follows:

- cost of cells £6000 – ICER £7414
- cost of cells £8000 – ICER £9700
- cost of cells £12,000 – ICER £14,272.

We tested a series of utility assumptions for those whose first repair was not successful but who decided not to have another. In our first analysis, we assumed that they had had some benefit, and had improved from a utility of 0.654 before the repair to 0.691 afterwards. NICE asked us to assess the effect of the following assumptions for utilities:

- Utility set to the same as failure (0.654) – ICER £15,634.
- Utility after failure set to same as success (0.817) – ICER £62,658. This assumption greatly increases utility gain among those who do not get good results after MF, and reduces the marginal QALY gains from ACI.
- Utility set to midpoint of success and failure (0.746) – ICER £27,123. This also reduces the marginal QALY gains from ACI as first procedure, because the larger proportion that does not do well after MF has their utility increased.

The Nawaz study provides very useful data on subgroups:

- Previous attempts at repair, such as MF – ICER £38,262. ACI is much less successful if the underlying bone has been damaged.
- Individuals without prior repair attempts – ICER £15,659.
- Kellgren–Lawrence grade 0 – no radiological sign of OA – ICER £15,618.
- Kellgren–Lawrence grade 1 – radiological signs of early OA – ICER £17,104.
- Kellgren–Lawrence grade 2 – ICER £20,096.
- Kellgren–Lawrence grade 3 – ICER £21,207.

In a SA, instead of relying on the Nawaz data alone, we tested the effect of pooling six ACI studies and found an ICER of £16,708.

Vericel provided details from an unpublished study in which patients with chondral defects were reported to have a baseline utility of 0.484. Using that baseline and their 3-year utility gain would give an ICER of £15,648. The baseline utility looks surprisingly low.

Strengths and limitations in evidence

We now have longer follow-up than was available for previous appraisals, and data from several new trials. The ACTIVE trial has data on some patients to 8 years and will eventually have 10 years of follow-up for all. The TIG/ACT trial has 5 years of follow-up. However, the two trials of MACI against MF had only 2 years of follow-up. There are few long-term MF studies.

Research needs

Autologous chondrocyte implantation is less successful among people with OA, but ICERs can be in the range usually considered acceptable. ACI may have a place in early OA with focal damage – research is needed in this group.

Conclusions

Caveats are necessary. There were more long-term studies of ACI than of MF. Using longer-term data than were available in the trials, MF comes out much less well. However, there are few long-term studies of MF, and extrapolation beyond observed data is subject to uncertainties. The evidence base is much stronger for ACI, but in older studies most patients had had previous attempts at repair. ACI is less successful after previous attempts at repair. Previous studies may therefore provide a pessimistic assessment. Most, but not all, studies suggest that ACI is more effective if used soon after the cartilage injury. A key conclusion is that ACI will give better results if used as first repair procedure.

In summary, the evidence base for ACI has improved since the last appraisal by NICE. In most analyses, the ICERs for ACI compared with MF appear to be within a range usually considered acceptable.

Study registration

This study is registered as PROSPERO CRD42014013083.

Funding

Funding for this study was provided by the Health Technology Assessment programme of the National Institute for Health Research.

Chapter 1 Background

History

The first appraisal of autologous chondrocyte implantation (ACI) was in 2000, after which the National Institute for Health and Care Excellence (NICE) issued Technology Appraisal (TA) Guidance 16,¹ which stated that:

1.1 Autologous cartilage transplantation is not currently recommended for routine primary treatment of articular cartilage defects of the knee joint in the NHS.

1.2 ACT should only be performed as part of a properly structured trial which wherever possible is randomized and adequately powered.

Reproduced with permission from NICE¹

This decision was made because there was then no evidence from randomised controlled trials (RCTs). The available evidence came from 17 case series of different interventions, and NICE concluded that:

Assessment of the evidence on clinical efficacy is confounded by a number of factors including variations in patient characteristics, concomitant surgery and use of multiple interventions. With one exception, all studies reported an improvement in patient status, usually over a follow-up period of less than 2 years.

These studies are summarised in the report by Jobanputra *et al.*² The studies lacked control groups, without which it is difficult to assess the effectiveness of procedures, relative to natural history or alternative treatments.

The guidance was reviewed in 2005, supported by a report by Clar *et al.*³ The guidance issued as TA89⁴ stated that:

Autologous chondrocyte implantation is not recommended for the treatment of articular cartilage defects of the knee joint except in the context of ongoing or new clinical studies that are designed to generate robust and relevant outcome data, including the measurement of health-related quality of life and long-term follow-up.

Reproduced with permission from NICE⁴

The terminology had changed. The initial term of 'autologous cartilage transplantation' had been replaced by 'autologous chondrocyte implantation' (ACI), which is more correct for two reasons. First, the small group of cells removed is multiplied before being put in, so transplantation is not correct because what goes back in is not what came out. Second, what is implanted is cells (chondrocytes) rather than cartilage, which takes time to develop.

The evidence base had improved by 2005, with four RCTs, two comparing ACI with mosaicplasty^{5,6} and two comparing it with microfracture (MF).^{7,8} The duration of follow-up was still short. At 2 years, there appeared to be little difference between ACI and mosaicplasty or MF. In the absence of long-term data, it was not possible to produce reliable costs per quality-adjusted life-year (QALY).

This report was written to support the third NICE appraisal of ACI in the knee.⁹

Chondral injuries

Articular cartilage covers the ends of the bones and the inner surface of the patella in the knee joint. It should not be confused with the meniscal cartilages that are cushions of cartilage between the bones – when people talk of ‘cartilage problems’ in the knee, they often mean the meniscal cartilage.

Normal hyaline cartilage is a rubber-like substance that is normally very smooth, promoting smooth, frictionless movements of the joints and also acting as a shock absorber. It is formed mainly of a protein called type II collagen. Under the articular cartilage are the bones of the knee: the femur in the thigh, the tibia below the knee and the patella or knee-cap.

Cartilage has no blood vessels and has very limited ability to repair itself. Epidemiological studies show a relationship between knee injury and later development of osteoarthritis (OA). In some people, this will lead in the long term to a need for a knee replacement with an artificial joint.

Loss of articular cartilage is referred to as a *chondral defect*, and loss of cartilage and bone as an *osteochondral defect*.

Cartilage damage can be caused directly from injury, by various types of arthritis, or spontaneously in a condition called osteochondritis dissecans (OCD). Cartilage damage may also arise because of knee instability or abnormal loading, for example secondary to a ligament injury¹⁰ or damaged meniscal cartilages.¹¹ Serious obesity may also affect knee cartilage.¹² Conversely, physical activity without injury may be protective.¹³

In young people the most common cause of hyaline cartilage damage is sporting injuries. Aroen *et al.*¹⁴ reported the causes of injury in patients having knee arthroscopy in Norway over a 6-month period. Injuries occurred in sport in 55%, in the home in 15%, at work in 12% and in road traffic accidents in 5%. In 13% the cause was unknown.

It should be noted that cartilage defects without any underlying bone involvement may not cause pain – there are no nerves in cartilage. The source of pain in knees with damaged cartilage is poorly understood but may come from many sources, including ligaments, the joint capsule and the underlying bone.¹⁵ Thus, results from series of symptomatic patients may not be entirely representative of all people with cartilage damage. The most common symptom is pain, with others being temporary locking of the knee in one position, and swelling. Pain and disability from symptomatic cartilage lesions has been shown to be as significant in magnitude as that from severe arthritis of the knee.¹⁶

The International Cartilage Repair Society (ICRS) has a scoring system for grading the severity of cartilage damage:¹⁷

- *Grade 1* Soft indentation and/or superficial cracks.
- *Grade 2* Small cracks or lesions extending down to under half of cartilage depth.
- *Grade 3* Deep cracks or gaps of over 50% of cartilage depth.
- *Grade 4* Cracks through the total thickness of cartilage down to the underlying bone.
- *Grade 5* Defects of the full thickness of cartilage involving the subchondral bone.

Grading has to be done by arthroscopic examination.

Interventions

Lavage and debridement

In lavage, the arthroscope (a sort of fiberoptic telescope) is inserted into the knee and saline is poured in through a cannula. This is usually done under general anaesthesia on a day-case basis. The saline washes out loose debris through the cannula. It is also thought to wash out compounds that cause inflammation.

Debridement is done under arthroscopic vision, and is the removal of damaged cartilage or bone.

Debridement and lavage are often done together.

The evidence for effectiveness is sparse and mixed. One three-armed RCT – of lavage alone, lavage plus debridement and a sham arm – reported no difference at 2 years.¹⁸ Another by Hubbard¹⁹ had methodological weaknesses, but reported that debridement and lavage was better than lavage alone. The NICE intervention procedures guidance (IPG230)²⁰ noted uncertainty about the efficacy of the procedure.

Autologous chondrocyte implantation

Cartilage cells are called chondrocytes. In ACI, a small piece of cartilage is removed from the knee, and the chondrocytes are grown in the laboratory until they number millions. They are then put onto the damaged area of articular cartilage as a patch. The hope is that this patch will repair the damaged area and form a new layer of natural articular cartilage, called *hyaline cartilage*.

Autologous chondrocyte implantation has been used since at least 1987,²¹ and the procedure has evolved over time. The Dutch Orthopaedic Association has provided a useful summary of developments.²² In the first generation of ACI, the cultured chondrocytes were placed in the defect, in liquid form, and then covered with a cap made from periosteum [autologous chondrocyte implantation–periosteal flap (ACI-P)]. This led to problems with pain in the immediate postoperative period and a need for further procedures to remove overgrowth in the graft as described in *Box 1*.

The second generation of ACI used a collagen cap [autologous chondrocyte implantation–collagen cap (ACI-C)] instead of the periosteal one, but still used cells in a liquid. Gomoll *et al.*²⁴ compared two cohorts, one that had a periosteal patch (ACI-P) and one that had a collagen cap (ACI-C). The reoperation rates were 26% and 5%, respectively. ACI-P is now little used in the UK, but is still used in the USA, where none of the membranes or scaffolds used in second-generation ACI has yet been approved by the Food and Drug Administration (FDA) except for in trials.²⁵

BOX 1 Clinical features of ACI-P

The periosteal patch was traditionally harvested via a 3- to 4-cm incision on the subcutaneous border of the proximal medial tibia. Careful dissection is performed to develop a plane between the periosteum (outer lining of bone) and overlying fat and fascia (outer lining of muscle). A slightly oversized patch is then harvested with a sharp surgical blade. This procedure takes approximately 30 minutes to perform and patients suffer from additional pain and swelling postoperatively. Potential complications include surgical site infection, and haematoma formation at the harvest site. If an infection does occur they are treated with a 1-week course of oral antibiotics.

The most common complications at site of implantation are graft overgrowth (hypertrophy) and scarring (arthrofibrosis) following this procedure. Overgrowth typically occurs between 3 and 6 months after the operation, and results from abrasion of the patch against internal structures in the knee. This can occur in up to 50% of cases, with a significant proportion requiring further keyhole surgery to debride ('shave off') the excess tissue from the surface of the patch.²³ Furthermore, suturing the patch may damage the native surrounding cartilage, as sutures are passed through normal healthy cartilage to ensure a watertight seal for the chondrocytes.

Contributed by Mr A Sprowson, orthopaedic surgeon.

In the third generation of ACI, the chondrocyte cells are loaded or embedded, or 'seeded', on to a porcine collagen membrane ACT-C (autologous chondrocyte transplantation seeded collagen membrane) or matrix [matrix-applied chondrocyte implantation (MACI®)], with a patch cut to fit. These patches can be implanted by a less-invasive form of surgery, by arthroscopy or mini-arthrotomy, requiring less surgical time than ACI-C.²⁶ (Arthrotomy = opening of a joint.) ChondroCelect cells (TiGenix, Leuven, Belgium) are now used in this way, with cells being loaded into the membrane by the surgeon.

The membrane used in MACI is composed of type I/III collagen, with a rough side wherein the chondrocytes are seeded and a smooth side that faces into the joint cavity.²⁶ The membrane is tough enough to be cut to shape or stitched in place, though it is more often glued in place.²⁶ The membrane is biodegradable. The term 'scaffold' is often used instead of membrane. However, the membrane needs careful handling to minimise chondrocyte death during implantation.²⁷

Another development, which can apply to both second- and third-generation ACI, has been that only selected chondrocytes are used – this is called characterised chondrocyte implantation (CCI). Cells that are most likely to produce hyaline cartilage with predominantly type II collagen, rather than a less resilient cartilage called fibrocartilage, which produces mainly type I collagen,²⁸ are identified during CCI using a panel of biomarkers, including collagen. TiGenix used six biomarkers and Genzyme (Sanofi) also used additional assays in CCI.²⁹

Table 1 summarises the generations of ACI. (Note: different authors use 'second generation' in different ways.) It is worth noting that graft hypertrophy can occur with second- and third-generation ACI. Niethammer *et al.*³⁰ reported graft hypertrophy on magnetic resonance imaging (MRI) in 11 of 44 patients who had MACI (with Novocart – TETEC AG, B. Braun, Reutlingen, Germany).

Harris *et al.*³¹ carried out a systematic review of failures and complications after ACI and reported that failure rates were higher with first-generation ACI-P than with second-generation ACI-C.

The Medical Services Advisory Committee (MSAC) report concluded that the ideal application of ACI would be in a full-thickness chondral defect surrounded by healthy cartilage in an otherwise healthy knee.³²

Microfracture

The main alternative method of repair is called 'microfracture', in which small holes are drilled through the surface of the bone in the area of damaged cartilage. This allows bleeding from the bone marrow, and the blood carries stem cells into the area where the damaged cartilage has been debrided. These cells form scar cartilage called *fibrocartilage*, composed of type I collagen. This is regarded as being inferior to hyaline cartilage, being less hardwearing and not expected to last as long.³³

MF may be combined with the insertion of a collagen membrane to cover the MF clot, known as *augmented MF*.

TABLE 1 The evolution of ACI

| Type of ACI | Method |
|----------------------------|--|
| First generation | ACI-P: Liquid suspension of cultured chondrocyte cells placed in the defect covered with a cap made from periosteum |
| Second generation | ACI-C: Liquid suspension of cells placed in the defect and covered with a collagen cap |
| Third generation | The cultured cells are seeded on to a membrane or 'scaffold' as in MACI (matrix-applied chondrocyte implantation) |
| Characterised chondrocytes | Not all chondrocytes are equally good at producing cartilage. Some are more 'chondrogenic' (cartilage producing) than others. The most useful can be selected and are known as 'characterised' |
| Fourth generation | Newer developments include the implantation not of cells that will form cartilage, but of tissue-engineered cartilage grown from autologous chondrocytes in collagen gel in the laboratory |

MF can be done arthroscopically (i.e. without opening the knee joint) and could be done at the same time as washing out a knee joint and stabilising loose tissue (debridement and lavage).

A search of the NICE website found no guidance on MF.

Mosaicplasty

Another method, which is now much less common, is mosaicplasty, sometimes called OATS (osteochondral autograft transfer system), which involves transplanting small sections of cartilage and underlying bone from a less-weight-bearing part of the knee into the damaged area. The pieces are in little cylinder shapes and, once transplanted, have an appearance not unlike a mosaic – hence the name. Mosaicplasty can be used only for small areas of damage (less than 4 cm²) because the transplanted sections have to come from elsewhere in the knee, usually the trochlea. (In some countries, allograft cadaver donor tissue is used, but this does not appear to happen in the UK.)

Mosaicplasty was reviewed by NICE through the Interventional Procedures Programme.³⁴ The guidance is reproduced in *Box 2*. It was dated ‘March 2006’ and so may now be out of date.

Mosaicplasty appears to be little used now. In the ACTIVE (Autologous Chondrocyte Transplantation/ Implantation Versus Existing Treatment) trial³⁵ (described in *Chapter 4*) of ACI versus standard methods, such as MF and mosaicplasty, few surgeons chose mosaicplasty.

Conservative management

Another option is no surgical treatment. Three case series^{36–38} reported high levels of return to activities after cartilage injuries after 14, 9 and 9 years, respectively. Messner and Maletius³⁶ reported a case series of young athletes (mean age 25 years, range 14–38 years) who had no treatment. Fourteen years later, most (21 out of 28) had returned to activity and 22 had excellent or good function.³⁶ However despite lack of symptoms, most showed radiological changes suggestive of early OA.

The British Association for Surgery of the Knee (BASK) UK Consensus recommends that all patients being considered for ACI should have had physical therapy first, as that may relieve symptoms.³⁹

BOX 2 National Institute for Health and Care Excellence Guidance. Reproduced with permission from NICE³⁴

Guidance

1.1 Current evidence suggests that there are no major safety concerns associated with mosaicplasty for knee cartilage defects. There is some evidence of short-term efficacy, but data on long-term efficacy are inadequate. In view of the uncertainties about the efficacy of the procedure, it should not be used without special arrangements for consent and audit or research.

1.2 Clinicians wishing to undertake mosaicplasty for knee cartilage defects should take the following actions.

- Inform the clinical governance leads in their Trusts.
- Ensure that patients understand the uncertainty about the procedure’s efficacy and the options for alternative treatments. They should provide them with clear written information. In addition, use of the Institute’s information for the public is recommended.
- Audit and review clinical outcomes of all patients having mosaicplasty for knee cartilage defects. The Institute may review the procedure upon publication of further evidence.

Decision problem

The scope from NICE for this appraisal mentioned three forms of ACI:

- The ChondroCelect ACI system from TiGenix, in which the cultured cells are combined with a biodegradable collagen I/III patch. This is a form of CCI. ChondroCelect received European marketing authorisation in October 2009.⁴⁰ It is marketed by Swedish Orphan Biovitrum (SoBi; Stockholm, Sweden). Production is being taken over by Pharmacell (Maastricht, The Netherlands).⁴¹
- The Matrix ACI system (MACI^{®*} – short for ‘matrix-applied characterised autologous cultured chondrocyte implant’) from Sanofi (Paris, France). The matrix refers to a collagen membrane with the chondrocytes. The Sanofi MACI[®] was approved in Europe in June 2013.⁴² This product is now being marketed by Aastrom Biosciences, who have changed their name to Vericel (Cambridge, MA, USA).
- ACI wherein the cells are cultured in hospital or research laboratories, such as the Robert Jones and Agnes Hunt Hospital (RJAH), Oswestry, termed ‘traditional ACI’ in the NICE scope. This appears to be the only UK NHS facility that currently cultures cells for use in ACI. Traditional ACI is used under hospital exemptions from the advanced therapy medicinal products regulations.

(*MACI is used both to refer to third-generation ACI and as a trade name; when referring to the trade name, we will use MACI[®].)

Autologous chondrocyte implantation is much more expensive than MF. The Australian Medical Service Advisory Committee estimated the cost of ACI to be about 10 times that of MF.³² In the NHS, the difference appears to be even greater, with list prices for cells alone being around £16,000.

The first decision to be made by NICE is whether or not ACI, in some or all of its forms, is clinically effective and cost-effective, and should now be used in routine NHS care. Both ChondroCelect and Vericel MACI[®] have marketing authorisations, with slightly different indications (*Box 3*).

It is not clear from the European Medicines Agency (EMA) website whether or not ChondroCelect is approved for lesions smaller than 1 cm².

The final scope for this appraisal did not consider sequencing of different technologies for the repair of cartilage defects, but the place of ACI in the treatment pathway needs to be examined. A second question is whether or not the much less expensive MF should be tried first, with ACI reserved for MF failures? Or are the best results with ACI achieved if it is the first treatment for chondral defects?

A third question is how soon cartilage defects should be treated. In a randomised trial of ACI versus MF, outcomes were better in those treated within 3 years of symptom onset compared with those with longer duration.⁴³

BOX 3 Licences for ChondroCelect and Verigen MACI[®]

ChondroCelect has a UK marketing authorisation for the ‘repair of single symptomatic cartilage defects of the femoral condyle of the knee (International Cartilage Repair Society [ICRS] grade III or IV) in adults’. The RCT that supported the marketing authorisation for ChondroCelect included patients with lesions of between 1 and 5 cm².⁴⁰

Vericel MACI[®] has a marketing authorisation for ‘the repair of symptomatic, full-thickness cartilage defects of the knee (grade III and IV of the Modified Outerbridge Scale) of 3–20 cm² in skeletally mature adult patients’.⁴²

Mithöfer *et al.*²⁸ have also reported better results with ACI sooner after injury, in football players. Harris *et al.*⁴⁴ also concluded that results were better in patients with shorter duration of symptoms and fewer prior procedures.

Thus, there may be a case for recommending earlier ACI.

Patient group

The patient group, as stated in the final scope from NICE, is 'adults with a symptomatic cartilage defect (chondral defect) but without advanced osteoarthritis'. Advanced OA is not defined in the scope. The chondral defects can be on the femur, tibia or patella. ACI is used in other joints, but such use is outwith the scope of this appraisal.

No age restriction is given in the scope from NICE, but in past trials, patients had a mean age of 32 years, range 16–49 years, with about 60% men. In most cases, the cartilage damage was due to injury, usually from sport.

Following a UK Cartilage Consensus meeting in March 2014, BASK produced a consensus document.³⁹ The points most relevant to this appraisal are summarised in *Box 4*.

Although the consensus recommends increased use of ACI, it would restrict it to people with symptoms and higher-grade lesions. As the statement recognises, some people may have symptoms relieved by physiotherapy. However, physiotherapy cannot repair chondral defects, so this group will still be at risk of progression to OA.

BOX 4 Consensus statement from BASK

The surgical management of symptomatic articular cartilage defects of the knee: consensus statement from UK knee surgeons

The statement notes variations in provision of repair of articular cartilage in the knee, and financial constraints on the more expensive treatment options.

The consensus relates to management of an isolated chondral lesion in a knee that is free of other defects, or in which these have been corrected. Key points include:

- Surgical treatment should be considered for symptomatic lesions of ICRS grade 3 or 4.
- MF leads to fibrocartilaginous scar tissue that has poorer biomechanical properties than normal hyaline cartilage, and this repair tissue degenerates. Short-term improvement in symptoms does not persist.
- Mosaicplasty can give good short-term results in small lesions, but longer-term results are poorer. It is not suitable for larger lesions or patellar defects.
- In small defects, less than 2 cm², MF, mosaicplasty and ACI may all be considered.
- For lesions > 2 cm², cell therapy (ACI) is the most effective treatment based on current evidence.
- Outcomes are poorer in smokers, patients with a BMI of > 30 kg/m², and those with a long duration of symptoms.
- When ACI is considered appropriate, it should be first-line treatment because results are poorer if it is used after failure of other procedures.
- Physical therapy may be effective in controlling symptoms and should be provided before surgery is considered.

BMI, body mass index.

Chapter 2 Clinical effectiveness

This chapter has three sections. First, we review some recent reviews on ACI and comparators, to give some general background. In this section, we provide information on most forms of ACI and how they compare with MF. We do this partly because the evidence on the technologies identified in the NICE scope is limited, both in terms of number of trials and duration of follow-up. This is a problem with evidence that is not unusual with non-pharmacological therapies:

- We need long-term follow-up.
- The technologies are evolving.
- By the time we get long-term follow-up from a study, the technology may have been superseded.

This is unlike the situation in drug appraisals where the drug molecule does not usually change over time.

Second, we give an account of two recent trials of MACI.

Third, we present the results from survival analysis based on a review of long-term observation studies of ACI and MF.

Systematic reviews

Inclusion criteria

Type of studies

We looked first for systematic reviews comparing relative effectiveness of ACI (any generation) and MF.

Type of participants

Adults with symptomatic articular cartilage defects.

Type of interventions

Autologous chondrocyte implantation for chondral defects in the knee only. All forms of ACI were considered.

Type of comparators

The main interest was MF but no restrictions were applied.

Type of outcomes

The outcomes of interest, as in the NICE scope, were pain and other symptoms, knee function including long-term function, rates of retreatment, activity levels, such as return to work or sport, avoidance of OA and knee replacement, adverse effects of treatment and health-related quality of life (QoL).

Searches for systematic reviews

Databases searched for systematic reviews published between 2004 and June 2014 were the Cochrane Database of Systematic Reviews, MEDLINE and EMBASE. The websites of the EMA, the US FDA and the Centre for Reviews and Dissemination (CRD) Health Technology Assessment (HTA) database were also searched for HTAs and other reports.

Detailed search strategies are provided in *Appendix 1*.

Study selection

Study selection was made independently by two reviewers. Discrepancies were resolved by discussion. There was no need for discussion with a third reviewer.

We selected recent reviews that provide comparative effectiveness data for ACI versus another comparator, but some reviews on other topics such as rehabilitation were also useful.

Data extraction strategy

Data were extracted by one reviewer and checked by a second using a standardised data extraction form. Discrepancies were resolved by discussion. There was no need for discussion with a third reviewer.

Quality assessment strategy

The quality of the reviews was assessed by one reviewer, and checked by a second reviewer. Any disagreements were resolved by consensus. There was no need for discussion with a third reviewer. The following quality criteria were used for assessing systematic reviews:

- inclusion criteria described
- details of literature search given (and adequate)
- study selection described (and adequate)
- data extraction described (and adequate)
- study quality assessment described (and adequate)
- study flow shown
- study characteristics of individual studies described
- quality of individual studies given
- results of individual studies shown
- statistical analysis appropriate.

Overall quality: high (≤ 1 of the criteria are not met)/medium (2–4 of the criteria are not met)/low (≥ 5 of the criteria are not met).

Methods of analysis/synthesis

Results were summarised narratively and in tables.

Results

The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram for systematic reviews is shown in *Appendix 2*.

The quality assessment of the reviews is reported in *Appendix 3*.

Twelve relevant systematic reviews were included. One by Vasiliadis *et al.*⁴⁵ was associated with a Cochrane review⁴⁶ but the former provides an update with more trials and is used here. The majority of reviews was rated as 'at least medium quality', with three reviews being rated as 'low quality',^{47–49} six reviews rated as 'medium quality' (Bekkers 2009,⁵⁰ Kon 2009,⁵¹ Magnussen 2008,⁵² Mithöfer 2009,⁵³ Nakamura 2009,⁵⁴ Negrin 2013⁵⁵) and three reviews rated as 'high quality' (Harris 2010,⁴⁴ Vasiliadis 2010,⁴⁵ Vavken 2010⁵⁶).

Table 52 in *Appendix 4* shows the primary intervention studies included in the reviews. Several reviews treated separate publications from the same study, or of subgroups of a study, as separate studies. We therefore checked the original studies and in the table we have grouped all reports from each study together. The tables describing the characteristics of the reviews also record publications from the same study.

The 12 reviews included 27 papers from 19 studies. Eleven of the studies were randomised trials (RCTs) and eight were comparative cohort studies or non-randomised/quasi-randomised trials. None of the primary studies was included in all of the reviews. Of the included primary studies, one compared collagen-based ACI with periosteum-based ACI, four compared ACI with MACI, one compared open with

arthroscopic ACI, three compared ACI with mosaicplasty, eight compared ACI with MF, and one each with bone marrow-derived mesenchymal stem cell therapy and with abrasion.

Characteristics of included reviews

Table 52 in Appendix 4 shows the characteristics of the included reviews. The reviews originated in various countries worldwide. None of the author teams appears to have had any specific conflicts of interest.

Objectives

Most studies sought to compare the effectiveness of ACI with that of other surgical treatments. Half of the reviews were very broad in their inclusion of comparators (Bekkers 2009,⁵⁰ Harris 2010,⁴⁴ Nakamura 2009,⁵⁴ Naveen 2012,⁴⁹ Vasiliadis 2010,⁴⁵ Vavken 2010⁵⁶), whereas others were more specific, for example comparing different generations of ACI (Goyal 2013⁴⁷) or focusing on MACI (Kon 2009⁵¹) and comparing with MF (Goyal 2013,⁴⁸ Negrin 2013⁵⁵) or osteochondral autograft transfer (OAT) (Magnussen 2008⁵²). One review focused on the effects of articular cartilage repair on athletic participation (Mithöfer 2009⁵³).

Inclusion criteria

Study design

The reviews included various types of study designs. They ranged from studies with very broad inclusion criteria [any type of primary study (Kon 2009⁵¹); RCT and prospective and retrospective studies with or without a control group (Mithöfer 2009⁵³); RCTs, prospective comparative studies and case series (Nakamura 2009⁵⁴) to studies including only level I and level II evidence/controlled trials or controlled prospective observational studies (Goyal 2013⁴⁷ and Goyal 2013,⁴⁸ Harris 2010,⁴⁴ Magnussen 2008,⁵² Negrin 2013⁵⁵) and studies including only RCTs or quasi-RCTs (Bekkers 2009,⁵⁰ Vasiliadis 2010,⁴⁵ Vavken 2010⁵⁶). Naveen 2012⁴⁹ stated that they would include only RCTs, but among the actual studies included were controlled clinical trials (CCTs) and comparative cohort studies. A few specified minimum follow-up times [6 months (Vavken 2010⁵⁶), 12 months (Harris 2010,⁴⁴ Magnussen 2008,⁵² Mithöfer 2009,⁵³ Negrin 2013⁵⁵)] and minimum number of participants (Magnussen 2008⁵²).

Participants

Inclusion criteria for participants were not given by all reviews. Some generally referred only to 'cartilage defects of the knee'; in others, the criteria were more specific, requiring full-thickness cartilage defects of the knee (Outerbridge grades III and IV: Harris 2010,⁴⁴ Magnussen 2008,⁵² Mithöfer 2009,⁵³ Negrin 2013,⁵⁵ Vasiliadis 2010⁴⁵) and in some cases also specifying anatomical location (femur, patella, trochlea: Mithöfer 2009,⁵³ Negrin 2013,⁵⁵ Vasiliadis 2010⁴⁵). An age range was specified only by Vasiliadis *et al.* 2010⁴⁵ (15–55 years).

Interventions

For most reviews, the index intervention was ACI. In two reviews the focus was on MACI/newer methods of ACI (Goyal 2013,⁴⁷ Kon 2009⁵¹). Magnussen 2008⁵² also includes OAT among the index interventions. In another review the index intervention was MF (Goyal 2013⁴⁸) and the authors reported outcomes only for MF, so the review is listed in the tables but will not be considered in the *Results* section. Comparators were not always explicitly stated, but included MF only (Goyal 2013,⁴⁸ Negrin 2013⁵⁵), MF or osteochondral autograft transplantation (Bekkers 2009⁵⁰), another ACI method (Goyal 2013⁴⁷), any cartilage repair technique or another generation of ACI or open versus arthroscopic ACI (Harris 2010⁴⁴), any other method (or placebo) (Magnussen 2008,⁵² Naveen 2012,⁴⁹ Vasiliadis 2010,⁴⁵ Vavken 2010⁵⁶) and any other method or no comparator (Kon 2009,⁵¹ Mithöfer 2009,⁵³ Nakamura 2009⁵⁴).

Outcomes

Often reviews did not explicitly specify outcome measures in their inclusion criteria. Whether specified or not, the focus was generally on (validated) clinical outcomes. Mithöfer *et al.* 2009⁵³ specifically focused on outcomes related to athletic activity. Many reviews also included information on the quality of the repair tissue and on complications.

Included studies

The reviews included between 3 and 13 comparative studies of individual populations relevant to this review (i.e. studies not including ACI or without a comparison group were not counted), with data on total numbers of patients ranging from around 200 to over 1000 participants. Individual study populations ranged between 19 and 231 participants.

As indicated above, 11 of the 19 comparative studies included were RCTs and eight were comparative cohort studies or non-randomised/quasi-randomised trials. Follow-up was between 6.5 months and 7.5 years (most reviews included studies with at least a year's follow-up). Many of the reviews commented on the quality of the studies, which overall was generally medium to low. Reasons included small sample sizes, inadequate durations of follow-up, lack of allocation concealment, and not enough information on method of randomisation, losses to follow-up and blinding of assessment scoring. Harris *et al.* 2010⁴⁴ reported that in their 13 included studies, quality was better in the later ones, but no studies were considered good or excellent – seven were scored as fair and six as poor. The origin of the included studies was generally not reported and only one review mentioned financial conflicts of interest of primary studies (Harris 2010⁴⁴).

Where reported, the mean age of participants was between 26.4 and 40.4 years, between 47% and 80% were men, and mean lesion size was between 1.9 and 6.4 cm². Lesion sites were mainly the femoral condyles, but sites such as the patella, trochlea and lateral tibia were also included. Both traumatic and non-traumatic lesions were included. Many of the participants had had previous surgery. Duration of symptoms before the intervention ranged between 1.5 and 10 years.

Box 5 shows the studies included in the reviews.

Gooding *et al.*²³ compared first-generation ACI-P with second-generation ACI-C, and found them similar in terms of repair quality, but with ACI-P requiring more subsequent procedures.²³ They concluded that ACI-C should be used and that ACI-P should be discontinued.

One trial by Bartlett *et al.* compared ACI-C and MACI (Verigen).⁵⁷ Both gave good results but MACI appeared slightly better, though most results were not statistically significant. (There were 44 patients in one group and 47 in the other.) The advantages of MACI were reported to be no need for suturing, a shorter procedure and a smaller incision. The proportions with good or excellent results were 72% with MACI and 59% with ACI-C.

Four studies compared ACI (mostly ACI-P) with MACI, one study compared open with arthroscopic ACI, three studies compared ACI with mosaicplasty, eight studies compared ACI-P with MF, and one study each with bone marrow-derived mesenchymal stem cell therapy and with abrasion. Clinical outcomes were measured using a wide range of different instruments. In some studies biopsies were also taken and histological outcomes reported.

Results and conclusions of reviews

The reviews generally agreed that studies were heterogeneous and had various quality limitations (as outlined above). The detailed results and the conclusions of the included reviews are provided in Appendix 6, Tables 54 and 55.

Clinical results

Improvements from clinical baseline scores were found regardless of treatment. One review suggested a small superiority of ACI (nine studies ACI-P, two ACI-C) compared with MF but not mosaicplasty (Harris 2010⁴⁴), but this review did not comment on the heterogeneity of results. Their forest plot⁴⁴ comparing MF and ACI showed three studies (Basad 2004⁸ and Basad 2010⁶⁴ with MACI®; Saris 2008⁶⁹ and Saris 2009⁷⁰ with ChondroCelect; Kon 2009,⁷³ MACI with Hyalograft) with better results with ACI, and one study

BOX 5 Autologous chondrocyte implantation: primary comparative studies in reviews**Autologous chondrocyte implantation–collagen cap vs. autologous chondrocyte implantation–periosteal flap***Randomised controlled trial*

Gooding 2006.²³

Autologous chondrocyte implantation vs. matrix-induced chondrocyte implantation*Randomised controlled trials*

Bartlett 2005⁵⁷ – MACI Verigen vs. ACI-C.

Zeifang 2010⁵⁸ – MACI vs. ACI-P.

Comparative cohort

Ergelet 2010⁵⁹ – MACI (BioSeed-C; BioTissue Technologies, Freiburg, Germany) vs. ACI-P.

Niemeyer 2008⁶⁰ – ACI-P vs. ACI-C vs. MACI (but each done by a different surgeon).

Open vs. arthroscopic autologous chondrocyte implantation*Comparative cohort/controlled clinical trial*

Ferruzzi 2008⁶¹ – MACI, open vs. arthroscopic.

Autologous chondrocyte implantation vs. mosaicplasty*Randomised controlled trial*

Bentley 2003⁵ – ACI-P.

Dozin 2005⁶² – ACI-P.

Controlled clinical trial

Horas 2000⁶³ – ACI-P (Described as RCT but inadequate randomisation method – alternation).

Horas 2003⁶ – ACI-P.

It is not clear whether the patients in Horas 2000⁶³ are included in Horas 2003.⁶

BOX 5 Autologous chondrocyte implantation: primary comparative studies in reviews (*continued*)**Autologous chondrocyte implantation vs. microfracture****Randomised controlled trial**

Basad 2004⁸ – This is presumably a preliminary report of the trial and patients reported in the first paper are expected to be included in the second report, Basad 2010.⁶⁴

Bachmann 2004⁶⁵ – This trial used MACI.

Crawford 2012⁶⁶ – MACI (Neocart – Histogenics, Waltham, MA, USA).

Knutsen 2004⁷ and Knutsen 2007⁶⁷ – ACI-P.

Lim 2012,⁶⁸ – ACI-P.

Saris 2008⁶⁹ – RCT ACI-P with CCI.

Saris 2009.⁷⁰

Vanlauwe 2011.⁴³

Van Assche 2009⁷¹ (Both Van Assche references involve the same subgroup of patients from the Saris RCT.)

Van Assche 2010.⁷²

Comparative cohort

Kon 2009⁷³ – MACI Hyalograft (Anika Therapeutics, Bedford, MA, USA).

Kon 2011⁷⁴ – MACI Hyalograft.

Minas 2009⁷⁵ – Case series on effect of previous MF.

Autologous chondrocyte implantation vs. bone marrow-derived mesenchymal stem cell**Comparative cohort**

Nejadnik 2010.⁷⁶

Autologous chondrocyte implantation vs. abrasionplasty**Randomised controlled trial**

Visna 2004⁷⁷ – MACI fibrin glue.

(Knutsen 2004 and 2007^{7,67} with ACI-P) reporting better results with MF. It was noted that the Knutsen^{7,67} results showed an advantage for MF at 2 years but not at 5 years. Harris *et al.*⁴⁴ concluded that MF showed an initial advantage that was then lost over time. They also concluded that there was a trend for ACI to show better outcomes than MF, but that a lack of long-term data meant that no definite verdict could be reached. Harris *et al.*⁴⁴ also commented on problems in interpretation due to the number of additional procedures undertaken in some studies, mainly meniscectomy and cruciate ligament repair.

Vakven *et al.* 2010⁵⁶ compared ACI (five ACI-P, one MACI, one fibrin glue) with mosaicplasty and MF, and were similarly cautious, mentioning 'a general trend for higher quality of repair tissue after ACI, suggesting better long-term results when compared to MF and osteochondral grafts' especially in higher-quality studies, but concluded that 'no clear recommendation can be deducted'.

Various reviews, including Vavken and Samartzis 2010,⁵⁶ questioned whether or not any small but significant differences seen in clinical outcomes were of real clinical importance. Significant differences between different generations of ACI were generally not seen. The delay in reaching maximal functional improvement (i.e. with respect to return to sports) may be slightly longer with ACI than with other interventions, but overall long-term durability may be greater with ACI.

Quality of repair tissue

The evidence suggested that ACI (all forms) may have a more durable repair tissue than MF (e.g. more hyaline-like cartilage).

Complications

Most notably, periosteum-based ACI was associated with a high rate of graft hypertrophy (over 20%) compared with only 3% with ACI-C (Harris 2010⁴⁴). Failure rates showed a reduction over the ACI generations: ACI-P 7.7%; ACI-C 1.5%; and 0.83% in all-arthroscopic second-generation ACI. Unplanned reoperation rates ranged from 27% with ACI-P to 1.4% in second-generation ACI. Harris *et al.*⁴⁴ found too few studies of third-generation ACI to report failure rates.

Modifying factors

Overall, outcomes tended to be better for younger patients [< 30 or < 35 years (age threshold varies among studies)], more active patients, patients with shorter symptom duration, and patients who had not had a previous failed surgical intervention. Results also tended to be better for smaller lesions overall, whereas ACI produced better results than MF in larger lesions (and its effect was largely independent of lesion size).

Recommendations for practice

Only five reviews made clear practice recommendations. Two of these (Vasiliadis 2010,⁴⁵ Vavken 2010⁵⁶) stated that the evidence was insufficient to recommend ACI over any other methods. The other three reviews agreed that MF was the first-line treatment for smaller lesions ($< 1\text{--}2\text{ cm}^2$) and that ACI was indicated for larger lesions ($> 2\text{ cm}^2$). The opinion about mosaicplasty was divided, with one review noting that its usefulness may be limited by donor site morbidity (Harris 2010⁴⁴).

The MSAC report³² also reviewed previous reviews and noted that most had been inconclusive, for reasons including:

- problems with the quality of the trials and other studies
- heterogeneity of patients recruited and of ACI and MACI techniques used
- variations in ages of recruits and size of defects
- variations in previous surgery
- multiple scoring systems and lack of standard outcomes
- safety data not reported as comprehensively.

Autologous chondrocyte implantation after previous microfracture

Microfracture is much less expensive than ACI, and is effective in the short term in most cases. It might therefore be suggested that MF should be tried first, and ACI used if it failed.

However, there is evidence that prior MF makes ACI less effective, because of a higher failure rate. This may be related to damage to the subchondral bone. Minas *et al.*⁷⁵ compared two cohorts of patients who had ACI-P, one group (111 patients) having had previous marrow stimulation procedures (MF, drilling or abrasion arthroplasty, all based on repair of the chondral defect by development of fibrocartilage from a blood clot) and the other (214) not. The groups were similar in age, duration and size of cartilage defect, duration of follow-up, concomitant procedures such as osteotomy or ligament repair, and size of repaired areas.

Failure was defined as persistence or recurrence of symptoms, or the need for a repeat procedure or knee replacement. The failure rate in those who had ACI as first procedure was 8% (17/214), but was 26% (29/111) in those who had had previous marrow procedures, and 20% in those who had had MF (but numbers small, 5/20).

Minas *et al.*⁷⁵ also report a subgroup of 15 patients who had more than one chondral defect (35 defects in total) about half of which had been treated by marrow stimulation and half not, with all then receiving ACI. The failure rate was 2 out of 18 in the previously untreated lesions and 16 out of 17 of the previously treated ones.

If ACI is less effective after prior MF, there are implications for the interpretation of results from some of the trials. For example, the Stanmore ACI trial⁷⁸ results were in patients who had had an average of 1.5 previous repair procedures. Only six patients had not had a previous repair procedure, so they could not compare results in those with/without previous surgery. Similarly, in a case series (Biant 2014⁷⁹) in patients with long-duration cartilage defects, those who had had previous procedures, such as MF, had 29% (21/72) failure of ACI compared with a 19% (6/32) failure rate in those having primary ACI. Failure was defined as requiring reoperation, somewhat stricter than in the Minas study.⁷⁵

One of the largest series of patients having ACI was reported by Nawaz *et al.* from Stanmore:⁸⁰ 1000 patients had ACI (519 with MACI, the rest ACI-C and some ACI-P) from 1998 to 2008.⁸⁰ In 827 patients with full follow-up data (mean follow-up 6.2 years), graft survival was 78% at 5 years and 51% at 10 years. Failure of the graft was 4.7 times as likely in the 34% who had had previous procedures (MF, mosaicplasty and drilling – numbers of each not given).

Pestka *et al.*⁸¹ reported a case series wherein 28 patients had MACI after previous MF and a matched 28 had MACI as first procedure. Failure was much more common in the previous MF group (7/28) than in the MACI as first procedure group (1/28).

There are two implications for this review. First, results seen in past trials wherein ACI was being used as a salvage procedure in patients with long-standing lesions and who had had previous procedures, may underestimate the benefits of ACI used as first procedure in patients with chondral defects of more recent origin. Second, a case could be made that ACI should be used as the primary procedure.

Other reviews

Mithöfer *et al.*⁸² carried out a systematic review of outcomes of MF, including 28 studies with 3122 patients, mean follow-up 41 months, with 1524 patients having follow-up of > 5 years. They noted good results in short-term functioning, but with need for further surgery increasing after 2 years, with rates of up to 31% by 5 years. Only five studies provided data beyond 5 years, of which one was a RCT and four were case series. At 6–7 years, most (67–86%) patients had improved knee functioning compared with baseline.

Several reviews examined factors that might predict success or failure. Behery *et al.*⁸³ reviewed 12 case series with 270 knees and found that none of age, gender, duration of symptoms and lesion size significantly predicted outcomes. They noted successful use of ACI in patients aged over 50 years in three studies. They concluded that the lack of association with lesion size made ACI preferable to MF in larger lesions. Another review from the same group⁸⁴ looked at factors that might influence the choice of repair method, and concluded that MF was less effective in larger lesions, when larger was defined (in different studies) as being greater than 2–4 cm².

Chalmers *et al.*⁸⁵ set out to systematically review activity-based outcomes [Tegner score, Lysholm score, KOOS, International Knee Documentation Committee (IKDC) score and the physical activity component of the Short Form questionnaire-36 items (SF-36)] after MF, ACI and mosaicplasty. They found only five studies that reported return to sporting activity. Return was faster after MF than ACI, but, beyond 2 years, activity scores deteriorated after MF but remained stable after ACI, though there was variation among sports. They noted the lack of long-term data on effects on later OA.

Mosaicplasty

Early results from the Stanmore trial (Bentley *et al.*⁵) showed good or excellent results in 88% after ACI-P or ACI-C compared with 69% after mosaicplasty, and the results at a minimum of 10 years' follow-up showed that repairs failed in 55% (23/42) of the mosaicplasty group and 17% (10/58) in the ACI group. For ACI, the patients in this trial were a difficult group, having a mean duration of symptoms of 7.2 years and an average of 1.5 previous procedures (excluding arthroscopy).

The Stanmore trial was omitted from the review by Harris *et al.* 2010,⁴⁴ which had only two studies of mosaicplasty, both favouring ACI, but with very wide confidence limits that overlapped with no difference.

The review by Vasiliadis *et al.*⁴⁵ identified three trials of mosaicplasty against ACI, two against ACI-P and one (Bentley *et al.*⁵) with both ACI-P and ACI-C. They reported that one trial (Horas 2003⁶) favoured mosaicplasty but another (Dozin 2005⁶²) found no difference. Vavken and Samartzis, reviewing the same studies, reported that the Horas trial⁶ showed no difference in clinical scores.⁵⁶

Bekkers *et al.*⁵⁰ concluded that single-plug mosaicplasty was the best option for small (less than 1- cm² osteochondral lesions).⁵⁰

The MSAC concluded that mosaicplasty should probably not be a comparator to MACI on the grounds of very low use in Australia.³²

Trials

Methods

Inclusion criteria

Type of studies

- Randomised controlled trials comparing second- and third-generation ACI and following patients for at least 2 years.
- Observational studies with at least 50 participants and follow-up of over 3 years were also considered, for results in routine care, adverse events (AEs) and costs.

Type of participants

- Adults with a symptomatic cartilage defect (chondral defect) but without advanced OA were included. The chondral defects can be on the femur, tibia or patella.
- The NICE scope did not report age restriction; however, we included studies comparing interventions of interest in patients aged 18 years and over.

Type of interventions

- ACI for chondral defects in the knee only. (ACI has also been used in shoulder, elbow, ankle and hip problems.) The forms of ACI considered were:
 - The ChondroCelect ACI, referred to by TiGenix as CCI.
 - The Matrix ACI system (MACI®) from Sanofi.
 - 'Traditional ACI' – the term used by NICE to describe ACI provided in the UK by hospitals that are using cells produced by non-commercial units, for their own use or for use in trials.

Type of comparators

- Microfracture is the main comparator. Mosaicplasty is now in limited use, for small defects only. Osteochondral grafts from cadavers can be used but are not to any significant volume in the UK and were not considered.

Type of outcomes

The outcomes considered, as also mentioned in the NICE scope, were as follows:

- pain
- knee function, including long-term function
- rates of retreatment
- activity levels, such as return to work or sport
- avoidance of OA and knee replacement
- adverse effects of treatment
- health-related QoL.

Table 2 summarises some of the outcomes used in ACI studies.

Howard *et al.*⁸⁶ carried out a high-quality systematic review to compare the various patient-reported outcome measures used in assessing the effects of ACI. They included 42 studies, grading the quality of studies with the Coleman methodology score. They concluded that the Lysholm and IKDC were the most responsive to change (i.e. showing larger effect sizes), but that IKDC and KOOS-Sports might reflect long-term outcomes better. They noted that the Cincinnati knee score also appeared satisfactory but based on few studies that there were several versions of this score, and many studies were excluded because the authors failed to state which version was used.

Exclusion criteria

We did not include trials of ACI-P in this section on the grounds that it had been replaced by third-generation ACI, but it should be noted that most long-term outcomes are from studies of first-generation ACI.

Search strategy

The databases that were searched for primary studies on clinical effectiveness published between 2010 and June 2014 were the Cochrane Central Register of Controlled Trials (CENTRAL), MEDLINE, EMBASE and the Web of Science.

TABLE 2 Outcomes used in cartilage repair studies

| Outcome measure | Variables included |
|----------------------------------|--|
| Lysholm score | Range of 0–100 (best), based on patient responses on eight aspects: pain, limping, locking, stair climbing, need for supports, instability, swelling and squatting |
| Tegner score | A level of activity measure from best 10, with ability to take part in competitive sports at a very high level, to worst 0, disabled |
| KOOS | Assesses pain, symptoms, ADL, sport and recreational activities, and knee-related QoL, with scores of 0 (worst) to 100 (best) |
| Cincinnati knee score | Based on symptoms (pain, swelling) and function (walking, climbing stairs, running) with a score of 0 (worst) to 10 (best). Variants include a sports rating from 0 to 100 points |
| ICRS score | This assesses quality of tissue repair rather than patient-reported outcomes. It could be argued that the quality of tissue repair might be useful for extrapolating from short-term histological results to long-term OA and need for knee replacement, but there is far from perfect correlation between symptoms and the degree of OA |
| IKDC score | Range 0 (worst) to 100 (best), based on function, symptoms, and range of motion. The version 'IKDC Subjective' is so-called because it is completed by patients |
| ADL, activities of daily living. | |

The inclusion lists of recent systematic reviews were also checked, and additional searches were done for ongoing or recently completed studies.

Auto-alerts in MEDLINE and EMBASE were run for the duration of the review to ensure that newly published studies were identified.

Details of search strategies are given in *Appendix 1*.

Identification of studies

Two independent reviewers screened titles and abstracts of the results were retrieved against the inclusion criteria. Those studies meeting the inclusion criteria were retrieved in full and checked for final inclusion by two reviewers independently. There was no need for discussions with a third reviewer.

Data extraction strategy

The data extraction template used by Harris *et al.*³¹ was used and adapted for this review. One reviewer extracted data, which was checked by a second reviewer.

Quality assessment strategy

The quality of the studies was assessed using the modified Coleman methodology score.³¹ There are 15 items in total, namely inclusion criteria, power, alpha error, sample size, randomisation, follow-up, patient analysis, blinding, similarity in treatment, treatment description, group comparability, outcome assessment, description of rehabilitation protocol, clinical effect measurement and number of patients to treat. A study could be rated as 'excellent' if the total score is between 85 and 100, rated as 'good' for scores between 70 and 84, rated as 'fair' with scores between 55 and 69, and, finally, categorised as 'poor' for scores of < 55.

The quality of the study was assessed by one reviewer and checked by a second reviewer.

Results

A total of 1672 records were retrieved by the searches. The title and abstracts were screened for inclusion and exclusion. Based on titles and abstracts, 104 records were considered possible inclusions and full texts of these were obtained. Out of 104 articles, two RCTs were included as definite inclusions and the

remaining 102 articles (which included the 12 systematic reviews included above) were excluded. The reasons for exclusion of 26 studies retained for final discussion by both reviewers is given in *Table 3*. (One of the excluded studies, reported in Saris *et al.* 2008⁶⁹ and 2009,⁷⁰ and Vanlauwe *et al.* 2011,⁴³ is described in the next chapter.)

MACI® versus microfracture

Two studies (Basad 2010,⁶⁴ Saris 2014¹⁰⁰) compared MACI against MF in patients with a symptomatic cartilage defect in the knee.

Basad et al. 2010

This RCT compared MACI, a third-generation ACI (then a Genzyme product) against MF in patients with symptomatic cartilage defects. Patients in the trial came from one centre (the principal author's clinic in Germany) between 2000 and 2005.

TABLE 3 Reason for exclusion of studies

| First author and year | Reason for exclusion |
|---------------------------------------|--|
| Bartlett 2005 ⁸⁷ | Technique includes bone graft |
| Bartlett 2005 ⁵⁷ | ACI (first generation) vs. MACI. 1-year FU |
| Benthien 2011 ⁸⁸ | Not a systematic review – no details of individual studies are given |
| Bentley 2003 ⁵ | ACI (first generation) |
| Bentley 2012 ⁷⁸ | ACI (first generation) |
| Cole 2011 ⁸⁹ | Not a form of ACI that we are including (CAIS) |
| Crawford 2012 ⁶⁶ | Not a form of ACI that we are including (NeoCart) |
| Dozin 2005 ⁶² | ACI-P |
| Ebert 2010 ⁹⁰ | Comparing rehabilitation approaches after MACI |
| Ebert 2012 ⁹¹ | Comparing rehabilitation approaches after MACI |
| Edwards 2013 ⁹² | Comparing rehabilitation approaches after MACI |
| Harris 2010 ⁹³ | Only includes one RCT that is not on ACI |
| Knutsen 2004 ⁷ | Old RCT of ACI-P |
| Knutsen 2007 ⁶⁷ | Five-year results from above trial. ACI-P |
| Lim 2012 ⁶⁸ | ACI-P |
| Panseri 2012 ⁹⁴ | Osteochondral defects |
| Rodriguez-Merchant 2012 ⁹⁵ | Short narrative review |
| Ruano-Ravina 2006 ⁹⁶ | Too old |
| Saris 2008 and 2009 ^{69,70} | ACI-P |
| Toonstra 2013 ⁹⁷ | Case series, only 20 patients, no controls |
| Trinh 2013 ⁹⁸ | About osteotomies, not ACI |
| United Healthcare 2013 ⁹⁹ | Not based on a systematic review |
| Van Assche 2009 ⁷¹ | ACI-P |
| Van Assche 2010 ⁷² | ACI-P |
| Ziefang 2009 ⁵⁸ | ACI-P vs. MACI and small numbers |

CAIS, cartilage autograft implantation system; FU, follow-up.

Quality assessment

Using the modified Coleman methodology score, the study scored a total of 45 suggesting that the quality of the study is poor, though this is partly due to failure to report items, so the study scored 0 points for those items. The enrolment rate was not reported, losing a maximum of 9 points. The power of the study (maximum score of 6 points) was not reported and it was not clear whether blinding of outcomes assessment (maximum score of 6 points) was done. There was no information available on effect size (maximum 6 points), relative risk (RR) reduction (maximum 6 points) and absolute risk reduction (maximum 6 points). There were some baseline differences between the two groups, so the study scored 6 points out of a possible 9 points. The study also lost points on the number of patients retained at the end of follow-up – 86.4% completed the 2-year follow-up period thereby scoring 4 points instead of a maximum 6 points.

Patient characteristics

Basad *et al.*⁶⁴ included 60 patients aged ≥ 18 and ≤ 50 years with a single symptomatic chondral lesion of femur or patella of size between 4 and 10 cm²; 40 received MACI and 20 MF. The mean ages of patients were 33 years in the MACI group, and 37.5 years in the MF group. The mean body mass index (BMI) of patients in the MACI was slightly lower than that in the MF group (25.3 kg/m² vs. 27.3 kg/m²). Previous surgery, if any, was not reported. Most defects in both groups were condylar (73% in MACI and 80% in MF), with the remaining lesions being in the patellar–trochlear region (28% in MACI and 20% in MF). Most patients were male (63% in MACI and 85% in MF). Patients in the MACI group had had symptoms for 2.2 years and those in the MF group for 2.5 years.

Details of intervention and comparators

Patients in the intervention group received MACI. The published paper states that the original protocol of the study had three interventions including two MACI groups and one MF group. In the two MACI groups, two different collagen matrices (supplied from two different manufacturing sites – name not reported) were used. The two matrices were considered identical in all aspects, so the two MACI groups were combined in the analysis.

Arthroscopy was done in all patients to assess their eligibility for the study (mainly isolated defect > 4 cm²). Patients in the MACI group had a sample from healthy cartilage sent for cell culture. Patients allocated to the MF group received treatment in one procedure. The MACI group returned 4–6 weeks later to have the chondrocyte-seeded collagen scaffold implanted into the defect, cell side down, facing the subchondral bone, and sealed with a thin layer of fibrin sealant.

Patients in both groups could also receive treatment for other concomitant lesions of cartilage or meniscus. All patients underwent a post-surgery rehabilitation programme. Those in the MF group received the rehabilitation programme recommended by Steadman *et al.*,¹⁰¹ which included 6 weeks of partial weight-bearing with 10-kg weight on crutches, continuous passive motion (CPV) and physiotherapy. After 6 weeks, patients were allowed to gradually progress into full weight-bearing.

The rehabilitation programme in the MACI group was slightly different. All patients had a plaster cast for 2 days after surgery in order to prevent graft delamination. Then, for the next 8 weeks, the programme included CPV, physiotherapy and partial weight-bearing with 10-kg weight on crutches.

All patients also received low-molecular-weight heparin each day during the partial weight-bearing phase to prevent deep-vein thrombosis (DVT).

Duration of follow-up

Patients were followed up for 2 years.

Outcomes

The primary outcome measures included the Tegner, Lysholm and ICRS scores. The Tegner score is related to activity levels of an individual, whereas the Lysholm score is related to pain, stability, gait and clinical

symptoms. The primary outcomes were measured at 8–12 weeks, 22–26 weeks and 50–54 weeks after surgery. One week after surgery, MRI scans were done in patients to see if there was delamination and graft hypertrophy. The efficacy population was defined as patients completing at least 6 months of follow-up, whereas completers were defined as those completing 2 years of follow-up. The definition of failure was not given.

Results

In total, 56 patients (39 in MACI and 17 in MF) completed at least 6 months of follow-up, and 48 patients (33 in MACI and 15 in MF) completed 2 years of follow-up. There was one early failure in the MF group but time was not reported. Two patients in the MF group (one pregnancy and one who had mosaicplasty) and one patient in the MACI group dropped out of the study.

There was improvement in the mean Lysholm score in both groups at year 1. The improvement in the MACI group persisted up to year 2 (52 at baseline, 95 at 12 months, 92 at 24 months), but it declined in the MF group after 12 months (55 at baseline, 81 at 12 months, 69 at 24 months). The improvement in Lysholm score from baseline to follow-up was statistically significant in both groups ($p < 0.0001$).

The improvement in the median Tegner score from baseline was greater in the MACI group than in the MF group. The Tegner score in the MACI group improved from level 2 to level 4 at 12 months, and remained at the same level at 24 months. The Tegner score in the MF group improved from level 2 to level 3 at 12 months, which was maintained at 24 months. The improvement from baseline to end of follow-up was statistically significant in both groups ($p < 0.0001$) but the improvement was statistically significantly greater in the MACI group than in the MF group ($p = 0.04$).

Some patients had issues with irritation during increased weight-bearing, treated with non-steroidal anti-inflammatory drugs (NSAIDs) and by returning to partial weight-bearing for a week. In the MACI group, one patient had persistent pain after 12 months and had arthroscopy, at which even and firmly regenerated cartilage repair was seen. The patient had persistent subchondral oedema. To relieve oedema, bone grafting was done.

Comments

The Basad group⁶⁴ has had long experience with ACI, so their results may be better than might be seen in routine care. Patients were treated with fairly short duration of symptoms, which may improve outcomes after ACI.

Saris et al. 2014

This trial¹⁰⁰ (SUMMIT, Superiority of Matrix-induced autologous chondrocyte implant versus Microfracture for Treatment of symptomatic articular cartilage defects) was a prospective, open-label, parallel-group, multicentre (16 European sites) RCT comparing Genzyme MACI (Genzyme, Europe) against MF.

Quality assessment

Using the modified Coleman methodology score, the study scored a total of 72 points, suggesting that the quality of the study is good. Information on blinding of outcomes assessment was not fully reported. There was no information on effect size, RR reduction and absolute risk reduction.

Patient characteristics

Patients aged between 18 and 55 years with one or more symptomatic cartilage defects, Outerbridge grade III or IV focal defects of size ≥ 3 cm² on medial or lateral femoral condyle and/or trochlea, and with a moderate to severe KOOS. There were 72 patients in each group. Most patients were male (62% in MACI, 67% in MF). Patients in the MACI group were slightly older than in the MF group (35 vs. 33 years). Mean BMIs were similar (26 kg/m²). 90% of patients in the MACI and almost 84% in the MF had undergone previous knee surgery. The most common prior procedures included diagnostic arthroscopy (50.3%), marrow stimulation techniques (in MACI group, MF 19%, drilling 11%), debridement of the

lesion (26.3%) and loose body removal (23.2%). Patients in the MACI group had had knee symptoms for longer than those in the MF group [mean of 5.8 years (range 0.05–28 years) vs. mean 3.7 years (range 0.1–15.4 years)]. The mean defect size of the lesions was similar across the groups (4.9 cm² in MACI and 4.7 cm² in MF). Most defects in both groups were on the medial femoral condyle (75% in MACI, 74% in MF), followed by the lateral femoral condyle (18% in MACI, 21% in MF) and trochlea (7% in MACI, 6% in MF). No tibial defects were reported.

Details of intervention and comparators

All patients underwent arthroscopy at baseline to examine their cartilage lesion and surrounding tissues. A small biopsy of cartilage (\approx 200 mg) was taken from a non-weight-bearing healthy area of the femoral condyle in all patients before randomisation, done using an interactive voice response system and computer-generated randomisation system. Those randomised to MF had it immediately. The technique recommended by Steadman *et al.*¹⁰¹ was followed, which included debridement and drilling multiple holes of centres 3–4 mm apart and 4 mm deep in the subchondral bone. Biopsies from patients receiving MF were preserved in case they later required MACI treatment. The MACI group had implantation of the cells 4–8 weeks after biopsy, by mini-arthrotomy. The MACI implant was trimmed to the size of the cartilage defect and implanted securely using a thin layer of fibrin sealant.

After surgery, both groups underwent the same rehabilitation programme, but individualised for patients. This was a four-phase programme recommended by Steadman *et al.*¹⁰¹

Duration of follow-up

Patients were followed up for 2 years. At the end of the follow-up, arthroscopy was performed to assess the condition of the knee.

Outcomes

The primary outcome measures were changes in KOOS for pain and function (sports and recreational activities subscore) subscales from baseline to year 2. Other outcome measures included histological evaluation of structural repair biopsy specimens, as measured by the microscopic ICRS-II overall assessment, and MRI assessment of the degree of defect fill, as measured by the scale of the whole-organ MRI score (WORMS: 0–25%, 26–50%, 51–75%, 76–100%).

In the study, response was defined as an at least 10-point improvement in KOOS for the pain and function subscales. Anyone not meeting both criteria was classed as a non-responder.

Failure was defined as any of the following:

- no improvement or worsening of the patient and physician global assessment
- less than 10% improvement in KOOS-pain
- a need for surgical re-intervention.

Those diagnosed as failures by physicians were further assessed by an independent treatment failure evaluation committee, who decided whether or not those cases were failures.

Results

A total of 144 patients was included in the study, 72 in each group. Ninety-five per cent (137/144) of patients completed the 2-year follow-up period. None of the patients in the MACI group discontinued treatment due to lack of efficacy, whereas three patients in the MF group discontinued study because of lack of efficacy.

The mean change in KOOS-pain from baseline to 2 years was significantly greater in the MACI group than in the MF group (45.5 vs. 35.5, difference between groups 11.76; $p = 0.001$). The change in the KOOS-function from baseline to 2 years was also significantly greater in the MACI group (46 vs. 36.1, difference

between groups 11.41; $p < 0.001$). Saris *et al.*¹⁰⁰ reported that the improvement in the KOOS-pain and pain score in the MACI over MF was observed at 36 weeks and maintained throughout the study period.

The proportion of responders was significantly greater in the MACI group than in the MF group (87.5% vs. 68.1%; $p = 0.016$) with more non-responders in the MF group (31.9% vs. 12.5%). Subgroup analyses found that more patients responded after MACI than after MF if patients had the following characteristics: male with a median age of < 34.5 years, only one lesion, lesions as a result of acute trauma, history of one previous surgery, symptoms for > 3 years (symptomatic response in those with under 3 years' duration: 82% with MACI and 69% with MF; over 3 years 92% and 67%) and size of lesions > 4 cm² and located on the medial femoral condyle. However, there were no statistically significant differences in response rates whether patients had or had not had previous cartilage surgery, as shown in *Table 4*.

In patients with larger lesions, ACI was reported to be more successful: 97% responders for MACI versus 77% for MF.

The improvements in other domains [activities of daily living (ADL), knee-related QoL, other symptoms] of the KOOS subscales were also statistically significantly greater with the MACI than with the MF. The mean differences between the two groups were:

- for the domain, ADL (mean change of 43.7 with MACI from baseline to 2 years; 33.2 with MF) at 2 years, estimated mean difference 12.01; $p < 0.001$
- for knee-related QoL (mean change of 37.4 with MACI from baseline, 30.1 with MF), estimated mean difference 8.98; $p = 0.029$
- for other symptoms (mean change of 35.4 with MACI from baseline, 27.8 with MF), estimated mean difference 11.61; $p < 0.001$.

At 2 years' follow-up, the modified Cincinnati knee score was significantly greater with MACI than with MF (1.05; $p = 0.002$). The IKDC score also showed favourable results for MACI (mean change from baseline with MACI 32.8 vs. MF 29.5); however, the difference between the two was not statistically significant ($p = 0.069$).

Comparison of treatment failure rates between treatment groups was not conducted because of the small number of failures: only two patients in the MF and none in the MACI group.

At 2 years' follow-up, 116 patients (60 in MACI, 56 in MF) underwent second-look arthroscopy and biopsy. There was good structural tissue repair with both treatments, and the repair was similar to the surrounding healthy cartilage. The mean ICRS-II overall assessment scores of the two treatments were similar (63.8 with MACI, 62.3 with MF, difference of 1.52; $p = 0.717$). The proportion of patients with overall assessment scores of normal or nearly normal (grade 1/2) was greater in the MACI group than in the MF group (76% vs. 60%).

At year 1 and year 2, 134 and 139 patients, respectively, underwent MRI evaluation. At year 1, the improvement was similar, but at year 2, more patients in the MACI group had a defect fill of $> 50\%$ of the defect depth than those in the MF group (83% vs. 77%).

TABLE 4 Response rates after prior cartilage procedures

| Prior cartilage surgery | MACI (%) | MF (%) |
|-------------------------|----------|--------|
| No surgery | 90 | 74.2 |
| 1 previous repair | 87 | 67.9 |
| > 1 previous repair | 84.2 | 53.9 |

More patients in the MF group complained of treatment-related AEs than in the MACI group (83.3% vs. 76.4%), the intensities of which were mild to moderate. The most commonly reported AE was arthralgia (57.6% overall – 51.4% MACI, 63.9% MF). Other events included back pain (11.1% MACI, 9.7% MF), joint swelling (9.7% MACI, 5.6% MF), joint effusion (6.9% MACI, 5.6% MF), pyrexia (5.6% MACI, 2.8% MF), cartilage injury (4.2% MACI, 12.5% MF), procedural pain (4.2% MACI, 5.6% MF) and ligament sprain (2.8% MACI, 5.6% MF). One patient (1.4%) in each group discontinued treatment due to AEs. More patients in the MF group had serious adverse events (SAEs) than in the MACI group (26.4% vs. 15.3%) such as treatment failure, cartilage injury and arthralgia.

Similar proportions of patients in the two groups underwent at least one subsequent surgical procedure (8.3% in MACI and 9.7% in MF). Two patients in the MACI group and none in the MF group underwent two subsequent surgical procedures. It has been reported that increasing age (not clear at what age) significantly decreased the likelihood of undergoing further procedures ($p = 0.038$).

Comments

Two factors will have reduced the chance of improvement: the long duration of symptoms before ACI (5.8 years) and the high proportion (37%) that had had previous surgery (not counting arthroscopy).

Summary of European Medicines Agency European Public Assessment Report

The EMA made a positive recommendation on MACI (manufactured by Genzyme Europe but then owned by Sanofi) on 25 April 2013. MACI has been recommended for the 'repair of symptomatic, full-thickness cartilage defects of the knee of 3–20 cm² in skeletally matured adult patients'.⁴² The product is available as an implantation matrix consisting of cultured chondrocyte cells on a membrane (500,000 to 1 million cells per square centimetre).

The clinical evidence on MACI came from the SUMMIT trial¹⁰⁰ (described above), which reported that MACI was better than MF in treating symptomatic cartilage defects of the knee, with size of the lesions ranging between 3 and 20 cm².

The EMA made a positive recommendation on ChondroCelect (TiGenix) on 25 June 2009.⁴⁰ ChondroCelect was recommended for the 'treatment of repair of single symptomatic cartilage defects of the femoral condyle of the knee (International Cartilage repair Society [ICRS] grade III or IV) in adults.'

The clinical evidence on ChondroCelect came from the TIG/ACT trial (described in detail in *Chapter 5 – Vanlauwe et al. 2011*⁴³), a phase III, randomised, multicentre trial comparing ChondroCelect with MF in patients with a single symptomatic cartilage lesion of the femoral condyles of the knee. At the time of appraisal, results from 12, 18 and 36 months were available, but we now have the 5-year results from Vanlauwe et al. 2011.⁴³

For discussion and conclusions on clinical effectiveness, see the end of *Chapter 5*.

Chapter 3 Survival analysis

Background

Following the second Appraisal Committee meeting, NICE requested additional work and further analyses from the Assessment Group.

In the assessment report, we focused on the second and third generations of ACI, on the assumption that the first generation, ACI-P, had been superseded by the later generations, because the new techniques were simpler and quicker and the use of periosteum was associated with the complexity of harvesting and ensuring a watertight cap, with overgrowth hypertrophy requiring reoperation and shaving of the graft, and the extra discomfort to patients from these procedures. The collagen cap is much easier to use but does come at an extra cost. The third generation of ACI in which the cells are seeded on to the collagen membrane is quicker still.

Because the second (ACI-C) and third (MACI) generations of ACI are fairly recent developments, we lack long-term data on their success rates. The TIG-ACT trial of ChondroCelect has 5-year follow-up,⁴³ but the SUMMIT trial¹⁰⁰ of MACI has so far published only 2-year results in full, with 36-month results in an abstract.

Therefore, NICE requested a review of all studies that provide long-term outcomes for ACI and MF, including both RCTs and observational studies, and all generations of ACI. In practice, if we define long term as more than 5 years, the ACI evidence comes from first-generation ACI; ACI-P.

There is some evidence to support extrapolating long-term outcomes after ACI-P to later generations. Gooding *et al.*²³ compared first-generation ACI-P with second-generation ACI-C and found them similar in terms of repair quality. There is no evidence that ACI-P has any advantages over ACI-C or MACI. (There was once a theory that the periosteal cap might promote chondrocyte function.) So, it seems reasonable to assume that data on longer-term outcomes of chondral defect repairs from studies of ACI-P can be extrapolated to survival of repairs after ACI-C and MACI. Niemeyer *et al.*¹⁰² compared ACI-P with ACI-C with 23 patients receiving each, matched for defect size and site, and age. Lysholm and IKDC scores were better with ACI-C (Lysholm, 63 versus 76, $p = 0.03$; IKDC 76 vs. 68, $p = 0.023$) but failures rates (defined as need for re-intervention) were the same by 10 years – 4 out of 23 in each group (17%).

Goyal *et al.*⁴⁷ carried out a meta-analysis to compare first-generation ACI with later generations, but found only three relevant studies, one of which was Gooding *et al.* 2006.²³ Niemeyer *et al.* 2014¹⁰² was not included. Goyal *et al.*⁴⁷ concluded that there was only weak evidence that ACI-C was any better than ACI-P because studies were up to only 2 years' duration and numbers were small. However, ACI-C was clearly no worse than ACI-P.

Autologous chondrocyte implantation–collagen cap was compared with MACI in one randomised trial from the Stanmore group. Bartlett *et al.*⁵⁷ randomised 91 patients to ACI-C or MACI. Follow-up was for only 1 year. The MACI group did better in symptoms, but the ACI-C group did better in cartilage quality. Despite randomisation, the ACI-C group had longer duration of symptoms (119 months vs. 88) and a higher proportion of previous failed procedures (20% vs. 4%), both of which are associated with poorer outcomes. However, the surgical team had longer experience of ACI-C than MACI.

In passing, it is worth noting the long duration of symptoms in many of the trials, and that this means that the results are likely to be worse than if ACI was used much sooner.

We therefore carried out a systematic review of long-term results of MF and ACI, defining long term as at least 5 years, not restricting study design, and assuming that the survival results of ACI-P could be extrapolated for modelling purposes to ACI-C.

Methods

- *Inclusions* Studies of any type of ACI that uses periosteal or collagen caps, or collagen matrices. Studies of MF, both traditional and capped [autologous matrix-induced chondrogenesis (AMIC)].
- *Exclusions* Studies of other forms of ACI, such as those using fibrin glue or synthetic caps or matrices not using collagen. Trials with fewer than 20 patients per arm. Observational studies with fewer than 40 patients. Studies of < 5 years' duration (even if a few patients have duration over 5 years). Trials or case series using drilling or abrasion methods. Studies where over 30% had significant concomitant surgery, such as tibial osteotomies, patellar realignment, or cruciate ligament repair were excluded.^{103,104} Minor concomitant surgery, such as partial meniscectomy, was allowed.

Search strategy

Searches, as shown in *Appendix 1*, were run in Ovid MEDLINE and Ovid EMBASE from 1997 to 15 May 2015. Thereafter, weekly auto-alerts, in MEDLINE and EMBASE, of these searches, were run until the end of 2015 to check for any new potential inclusions.

The searches retrieved 2907 documents; after removing duplicates and animal studies, 1833 records remained and the title and abstracts were screened by two authors for inclusions. The full text of 69 articles was checked: 26 articles (21 studies) were included and 43 articles were excluded.

Quality assessment used the National Institutes of Health (NIH) checklist for observational studies.¹⁰⁵

Results

A broad search with no restriction on designs (to capture case series) retrieved an initial 1833 studies, of which 67 were possible inclusions, based on abstracts.

Table 5 shows the included studies. Not all were used in survival analysis. Excluded studies are listed in *Appendix 7*. Details of all the studies used are provided in *Appendix 8*.

Comparing results of different studies is not straightforward, because a number of factors influence the results, including:

- Previous attempts at repair – these reduce the chance of success. Most of the older studies had patients who had had unsuccessful previous surgery.
- Size of defect. Large lesions do not do well with MF.
- Site of defect. For example, ACI appears to be less successful in trochlear lesions. Some studies exclude trochlear lesions (Knutsen 2007⁶⁷).
- Duration of chondral defect.
- Surgical experience and learning curves.
- Length of follow-up and losses to follow-up.
- Age.
- BMI.
- Activity levels after repair. Some studies are in elite sportsmen and women, who may put great demands on the repair. Some patients may go back to activity too early.
- Concomitant surgery, or lack of it. For example, there was only one concomitant osteotomy in the Moseley series,¹¹⁵ but some patients had misalignment, which, left uncorrected, increased the failure rate.
- Outcome measures used – reoperation or symptom scores.
- Registry requirements/criteria.

The strong adverse relationship between prior attempts at repair and failure mean that most of the older studies will give a misleadingly pessimistic picture if applied to ACI carried out in people with recent-onset

TABLE 5 Included studies

| Author | Brief description | Quality assessment and whether or not used in survival analysis |
|------------------------------|---|---|
| Asik 2008 ¹⁰⁶ | Assessment of a series of 90 patients after MF. Those having other procedures as well were excluded. Mean FU; 68 months. No failure data reported. Better results in those treated sooner (< 12 months) after injury, in younger (< 35 years) people, defects < 2 cm ² and BMI of < 25 kg/m ² . No data on previous procedures | Good, no |
| Bentley 2012 ⁷⁸ | Long-term (minimum 10 years, range 10–12 years) results of the 58 patients in the ACI arm from RCT vs. mosaicplasty (Bentley <i>et al.</i> ⁵). Loss to FU 9%, censored at last visit. ACI-P or ACI-C. Long duration of injury before ACI (mean 7 years, range 1–20 years) and 94% had had previous surgery such as MF. 175 failure rate – graft failure or reoperation. Mean defect size 4.6 cm ² , range 1–10 cm ² | Good, subsumed into the Nawaz study ⁸⁰ in survival analysis, but see <i>Discussion</i> |
| Beris 2012 ¹⁰⁷ | Case series of 42 patients (45 knees) after ACI-P. Mean defect size 5.3 cm ² , range 2–12 cm ² . Mean duration 28 months. No data on prior surgery, loss to FU or failures | Poor, no |
| Bhosale 2009 ¹⁰⁸ | Cohort study of first 80 consecutive patients having ACI-P 1996–2002. ACI-only. 87.5% previous surgery. Duration of defect not reported. Failures not reported. Mean defect size 4.1 cm ² , IQR 3–6. If success at 15 months, sustained for up to 8 years | Good, no |
| Biant 2014 ⁷⁹ | Case series 104 patients after ACI-P (19) or ACI-C (85) in 1998–2001, followed up for at least 10 years. Duration defect 7.8 years, size 4.8 cm ² , range 1–25 cm ² . Loss to FU 4%. Previous surgery in 70% and they had poorer results. Failures in 26% at mean FU 5.7 years (all by 8 years), defined as revision of repair or arthroplasty. Results in non-failures good or excellent in 88% | Good, subsumed under Nawaz study ⁸⁰ in survival analysis, but see <i>Discussion</i> |
| Browne 2005 ¹⁰⁹ | Case series. Clinical outcome of ACI at 5 years in 87 subjects. Defect size 4.9 cm ² , duration not reported. Previous failed surgery 70%. In 36%, the ACI was the first performed by surgeon. Of the 87, 62 (70%) improved, six no change, 19 were worse | Fair, no |
| Gomoll 2014 ¹¹⁰ | ACI in the patella only. Case series of 110. 8% failures, defined as graft failure with pain, requiring revision surgery | Fair, no |
| Jungmann 2012 ¹¹¹ | Cohort study of predictors of failure 2–12 years after ACI: 26% ACI-P, 57% ACI-C, 17% BioSeed. Failure defined as need for revision surgery. <i>N</i> = 413. Prior repairs 70%; with more than one repair, 16%. More than one prior repair increased risk of failure fourfold. ACI-P doubled risk of failure vs. ACI-C. No association of age or BMI with failure but only 85 had a BMI of > 30 kg/m ² . Duration of defect not reported | Fair, used |
| Knutsen 2007 ⁶⁷ | This RCT of ACI vs. MF showed no difference. <i>N</i> = 80. 5-year FU. 23% failure in both arms, defined as need for revision. 93% had had prior repairs. Median duration of defect 3 years. At baseline, people with OA excluded but by 5 years, 34% had Kellgren–Lawrence grade 2 or worse in their late thirties. Younger patients did better (< 30 years) | Fair, used |
| Krych 2012 ¹¹² | Case series of 48 MF patients from comparative study of mosaicplasty and MF | Not used in survival analyses |

continued

TABLE 5 Included studies (continued)

| Author | Brief description | Quality assessment and whether or not used in survival analysis |
|--|--|---|
| Layton 2015 ^{113,114} | Layton <i>et al.</i> report results of MF in a very large observational study of 3498 patients in the USA. The data were obtained from an administrative claims database. The study has not yet been published in full, but is available as an ISPOR abstract. The authors have provided a copy of the full poster | Good, yes |
| Moseley 2010 ¹¹⁵ | Registry-based case series of 72 patients from 35 centres, 24 of which entered only one patient. In 29%, the ACI was the first procedure done by the surgeon. Mean defect size 5.2 cm ² . 36% had had previous attempts at repair. Duration of defect not reported. 21% had concomitant surgery but mostly minor with only one osteotomy. Failure, defined as need for reoperation, occurred in 17%. At 6–10 years, 69% of patients had good results | Fair to good, yes |
| Nawaz 2014 ⁹⁰ Incorporates Rogers 2010 ¹¹⁶ with the Briggs series Also incorporates data from the ACI arm of the Bentley 2012 trial ⁷⁸ and the patients in the Biant 2014 study ⁷⁹ | Long-term study ACI in 827 patients with mean FU 6.2 years (range 2–12 years) after ACI-P or ACI-C) in 37% or MACI (63%). 499 had reached 5 years of FU and 366 had reached 8 or more years, making it one of the most useful studies. Mean defect size 4.1 cm ² , range 0.6–20.8 cm ² . 34% had had previous cartilage repair surgery, and they were five times as likely to fail ACI. Failure was defined as need for further surgery, graft delamination (MRI or arthroscopic) or symptom scores close to, or worse than, pre operation. Early OA was associated with poorer outcomes – HR for failure 2.1 for Kellgren–Lawrence grade 1; 3.5 for grade 2; and 3.8 for grade 3. Defect size did not affect failure risk | Good, yes |
| Niemeyer 2014 ¹¹⁷ | Case series of ACI-P. <i>N</i> = 86 but 16 lost to FU. Duration of defect 'several years', mean size 6.5 cm ² . 34% prior repair attempts. Some concomitant surgery but mainly partial meniscectomy. 29% had further surgery but not all related | Good but for 19% dropout rate |
| Niemeyer 2014 ¹⁰² | Matched-pair comparison of outcomes of ACI-P vs. ACI-C. 23 per group and FU at least 10 years. Same failure rate: 4 (17%) in each group required further surgery including TKR. ACI-C better on Lysholm and IKDC scores. But small study | Good, no |
| Peterson 2010. ¹¹⁸ Includes patients reported in Peterson 2002 ¹¹⁹ with chondral injury and 26 of those in Peterson 2003 ¹²⁰ | Long-term FU of the Gothenburg patients of Brittberg <i>et al.</i> , ²¹ who had had ACI-P at least 10 years before (but range given as 9.3–20.7 years) with mean FU 12.8 years; 341 questionnaires sent out and 224 replies (65%) despite many having moved. Lysholm, KOOS, etc., plus question about whether or not they would have again – over 90% would; 74% reported better or same, 26% worse. No data on failures requiring reoperation provided. Neither age nor size of lesion affected outcome. Size 5.3 cm ² mean lesion size, but some had more than one lesion. So majority had good result 10–20 years later | Fair, not used |
| Salzmann 2013 ¹²¹ | Reoperative characteristic after MF of knee cartilage lesions in 454 patients. Retrospective chart review Mean FU duration: ● Failure subjects 5.0 ± 2.1 ● Non-failure subjects 4.4 ± 1.9 | Not used in SA |

TABLE 5 Included studies (continued)

| Author | Brief description | Quality assessment and whether or not used in survival analysis |
|---|---|---|
| Shive 2015 ¹²² (also Frappier 2014 ¹²³) | This study reports a multicentre RCT of MF with BST-CarGel® (Piramal Life Sciences, Mumbai, India) vs. MF alone at 5 years. There was no ACI arm. The trial was of a form of enhanced MF, with a chitosan framework to stabilise the blood clot. At 5 years, there was no difference in clinical outcomes, but the quality of the cartilage filling was better with CarGel. Whether this would result in later clinical benefits from a longer-lasting repair is not yet known. A cost-effectiveness analysis (Frappier 2014 ¹²³) making assumptions on failure and fill, reported that BST-CarGel could be cost-saving | Fair, not used |
| Solheim 2014 ¹²⁴ (incorporates patients from Solheim 2010 ¹²⁵) | MF treatment of single or multiple articular cartilage defects of the knee: a 5-year median FU of 110 patients | Not used |
| Steadman 2003 ¹²⁶ | FU of cohort of 72 patients after MF for traumatic chondral defects of the knee. Average FU 11 years, range 7–17 years. Duration of defect mean 3 years, range 9 months to 7 years. Mean size 2.8 cm ² . Only two failures, which were excluded from study. Of 71 followed for 7 years, 59 had improved, 11 were the same and 1 was worse. Unusually low failure rate. Patients were selected from a larger (302) group by excluding those with other lesions, degenerative change and concomitant surgery. Note quite small size of lesion and inclusion of children | Fair, used |
| Vanlauwe 2011 ⁴³ | 5-year results from the TIG-ACT trial of ChondroCelect vs. MF, dealt with earlier in assessment report | Good, used in SA |

FU, follow-up; HR, hazard ratio; IQR, interquartile range; ISPOR, International Society for Pharmacoeconomics and Outcomes Research; SA, sensitivity analysis; TKR, total knee replacement.

defects, for which ACI is the first procedure. Nevertheless, some studies in which ACI was a last-resort salvage procedure reported good results in many patients.

Results of autologous chondrocyte implantation

The most useful study is that from the Stanmore group, by Nawaz *et al.*,⁸⁰ because this study is the largest, is of good quality and reports UK practice, albeit from a centre of excellence. The Nawaz paper⁸⁰ reports results in 827 patients, which allows for very useful subgroup analysis. Mean age at baseline was 34 years, range 14–56 years. Radiographs were taken and assessed for degenerative change according to the Kellgren–Lawrence grading. The ACI procedures were carried out from 1998 to 2008, and all patients were assessed in 2010, allowing a Kaplan–Meier (KM) survival curve to be constructed for over 10 years. Thirty-four per cent (282) of the patients had had previous repair attempts, such as MF, and they had much poorer graft survival by 10 years: under 25% compared with 75% in those who had had no previous procedures. The recruitment period spanned the generations from ACI-P to ACI-C and on to MACI. There was no difference in survival time between ACI-C/ACI-P and MACI.

Patients with Kellgren–Lawrence grades 2 and 3 had only 25% graft survival by 10 years. Those with grade 1 fared better initially, but by 10 years were catching up on the grades 2 and 3. Those with no degenerative change did much better, with about 70% graft survival at 10 years.

In summary, results of ACI are poorer in:

- patients who have had previous attempts at repair
- those with early OA as reflected in Kellgren–Lawrence grades 1–3.

Combining these led to very poor results – ACI in a patient with previous repair and Kellgren–Lawrence grade 3 had little chance of survival at 10 years.

Size of lesion did not affect survival.

Results of microfracture

The long-term evidence on MF was sparse. We note a comment by Salzmann *et al.*¹²¹ that:

The general body of literature concentrating on the clinical outcome following microfracture at the knee joint is surprisingly light when compared with its clinical popularity.

This applies particularly to studies reporting outcomes beyond 5 years. There are few of these so we relaxed our exclusion criterion of a minimum of 40 patients in observational studies.

Gobbi *et al.*,¹²⁷ in a series of 61 patients followed for 15 years, reported good results at short-term follow-up, but that deterioration could be expected after 2–5 years. Their failure rate, defined as need for reoperation, was only 7 patients (11%), but 40% showed osteoarthritic changes.

Gudas *et al.*¹²⁸ reported outcomes in the MF arm of a trial against mosaicplasty. In 29 patients, 11 (34%) had failed (required reoperation) by 10 years. Most failures occurred by 40 months. Defects averaged 2.8 cm².

Solheim *et al.*¹²⁵ followed up 110 patients for 10–14 years after MF, and reported failure, defined as need for further surgery in 39% and poor results (Lysholm score of 64 or less or needing knee replacement) in 46%. They commented that although outcomes score improved after MF, normal knee function was usually not achieved.

Steadman *et al.*¹²⁶ reported better results in a series of 72 patients followed for an average of 11 years, range 7–17 years. Their average defect size was 2.8 cm² and mean duration of injury was 3 years (range 9 months to 7 years). Two failures were excluded, as were patients having concomitant surgery, or who had OA, and those over 45 years – so the 72 patients are a subset of 302. At year 7, 59 (80%) had improved by having less pain. The good results may reflect the good prognostic indicators.

By far the largest MF study has not yet been reported in full, and unfortunately provides follow-up data only up to 5 years. Layton *et al.*¹¹³ from Quorum Consulting reported results in 3498 patients using data from an American claims database, published as an abstract from the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) conference¹¹³ but with greater detail available in the poster that is on the Quorum website.¹¹⁴

Not all of the patients reported are relevant to this report, because they included 351 Medicare patients with average age 73 years. We excluded them. And even the ‘commercial’ group is older – at mean age of 47 years – than most patients being considered for ACI. However, they do provide a good guide to the success of MF in routine care and have impressive numbers. Layton *et al.*¹¹⁴ reported failure rates (further surgery) of 9% within 1 year, 18% at 3 years, and 32% by 5 years. Data on analgesic consumption suggests that others did not have further surgery, but needed opiate or other analgesia.

The future of MF has been reviewed by Bert,¹²⁹ who points out that the landmark studies of MF, such as that by Steadman *et al.*,¹²⁶ did not have control groups of debridement alone. Bert¹²⁹ cites the 2013 study by Gudas *et al.*¹³⁰ as the only trial in which MF was compared with debridement alone, and which showed no difference. Unfortunately, the Gudas *et al.* trial¹²⁴ was quite small (34 patients per arm), had only 3 years' follow-up, and would not score well on the Cochrane risk of bias checklist. Bert¹²⁹ argues that debridement alone will give as good results as MF, but without damaging the underlying bone, which would reduce the likelihood of success with later ACI.

Time-to-failure studies

Caveats

When considering survival curves extrapolated beyond the observed data, it should be borne in mind that the extrapolation assumes that the curve based on the observed data will continue. When using parametric fits for extrapolation (any fit irrespective of the equation that describes it), the usual option is to select what is considered to be the best fit to the observed data. Unfortunately, there is no universally applicable method to determine the best fit and opinions may differ. With some data, most well-fitting models will produce similar extrapolations; however, with ACI data this has not been the case. Selecting several plausible models that produce different extrapolations should bracket what can be argued to be the best estimate of behaviour beyond the observed data.

In most instances it was possible to reconstruct individual patient data (IPD) using published data; in the case of degenerative change reported by Nawaz 2014,⁸⁰ the only way to get parametric curve parameters was using a digitised KM (rather than IPD) with the non-linear regression Stata® (StataCorp LP, College Station, TX, USA) command specifying candidate parameters. Since in this instance it was not possible to get a meaningful covariance matrix these particular results could only be used in deterministic cost-effectiveness analysis and not for probabilistic sensitivity analyses (PSAs). Occasionally with some other parametric models no confidence interval (CI) could be derived (probably due to lack of numbers for patients and for events) and so uncertainty was not quantifiable (see *Appendix 9*); again in these instances for cost effectiveness estimates only deterministic analysis was possible.

Methods

Published KM graphs were digitised and IPD reconstructed using the Guyot *et al.* method.¹³¹ Where published graphs had the appearance of KM plots but authors did not specify their method, it was assumed to be KM. Where plots were presented as scatter graphs rather than stepped lines it was assumed data points represented the top, rather than bottom or midpoint, of a stepped fall in survival.

Parametric models of time to failure were used to explore failure rate beyond the observed data. In the absence of patient numbers this was done by least squares non-linear regression. Where patient numbers allowed use of the Guyot method to reconstruct IPD, the models were implemented in Stata (version 12) with the 'streg' command and or using the 'stgenreg' package of Crowther and Lambert 2013.¹³² Standard models (exponential, Weibull, log-logistic, log-normal, Gompertz and gamma) were explored together with additional models for increasing hazard through time, either after an initial phase of decreasing hazard (bathtub model) or with linear increase in hazard (Rayleigh models). CIs (95%) were estimated with the delta method. The bathtub model was investigated because it has previously been found useful for modelling failure rates after total hip replacement.¹³³

Linearly increasing hazard models were tried because MF failure rate during 12 years of follow-up of patients with OA (Bae *et al.*,¹³⁴ an excluded study for this report, because patients had OA and were older)

was found to be best fit with such models. They were therefore judged to be worth investigating. Linear hazard models were used with one or two parameters in which survival is described as:

$$A] S = \exp(-\lambda \times \text{time}^2) \quad (1)$$

$$B] S = \exp(-(\lambda_0 \times \text{time} + \lambda_1 \times \text{time}^2)) \quad (2)$$

($\lambda_0 \geq \text{zero}$; when λ_0 is zero, S conforms to equation 1.)

Model fit was judged using information criteria and by visual inspection of cumulative hazard plots¹³⁵ and KM plots. One study provided Cox multivariate regression analysis of patient subgroups. The hazard ratios (HRs) from these analyses were used to estimate subgroup failure rates using two methods: (1) A log-normal model hazard was calculated for the baseline subgroup and multiplied by the appropriate HR for each of the other subgroups. The time to failure for these subgroups was estimated from $\exp(-\text{cumulative hazard})$. (2) A Weibull or linearly increasing hazard model survival for the subgroups was estimated from: $\exp(\ln(\text{baseline subgroup survival}) \times \text{HR})$.

Log-normal model hazard was calculated from:

$$\text{Hazard} = A/B \quad (3)$$

where:

$$A = C \times D \quad (4)$$

$$C = \exp(-0.5 \times ((\ln(\text{time}) - \mu)/\sigma)^2) \quad (5)$$

$$D = (\text{time} \times \sigma \times \sqrt{2} \times \sqrt{\pi})^{-1} \quad (6)$$

$$B = 1 - \Phi((\ln(\text{time}) - \mu)/\sigma) \quad (7)$$

where Φ is the standard normal distribution.

Parametric models and KM plots are presented in *Appendix 9*. The remit for this report was to exclude studies with less than 5 years of time-to-event data; therefore, except in exceptional circumstances, such studies were excluded.

Description of time-to-failure data

Seven relevant published studies^{43,67,80,111,115,117,136} were identified that presented KM plots extending to at least 5 years. Estimates of IPD reconstructed from such plots are best served (i.e. likely to be more accurate) when the total number of events, and of patients, are reported; when accompanied by a risk table indicating the number of participants remaining at risk at multiple time intervals; and when the graphical display is of sufficient quality to unambiguously identify the times at which events occurred. Extrapolation of parametric models' fit to such IPD are more likely to be reliable for more mature data, that is when follow-up is sufficient that the number of events has reduced the probability of survival at the end of the plot to a low value. A rough estimate of maturity is whether or not median survival has been reached.

Some of the included studies^{80,136} aggregated event times to yearly intervals. This may be because precise times of failure were not recorded. The impact of this on subsequent use of data is difficult to gauge. Risk tables were rarely presented, and in some studies reporting subgroup analyses the number of patients, as well as the number of events, for some subgroups was not reported. Median survival was reached in only one study. One study¹³⁶ presented KM analyses in scatter plots rather than a more conventional stepped plot. In this study, it was difficult to be certain whether data points represented the top, bottom or midpoint of a step in survival. These characteristics are summarised in *Table 6*.

TABLE 6 Time-to-event data presented in relevant studies

| Item/study ^a | Knutsen 2007 ⁶⁷ | Minas 2014 ¹³⁶ | Moseley 2010 ¹¹⁵ | Nawaz 2014 ⁸⁰ | Niemeyer 2014 ¹¹⁷ | Vanlauwe 2011 ⁴³ |
|-------------------------|----------------------------|---------------------------|-----------------------------|--------------------------|------------------------------|-----------------------------|
| Patient no. | Yes | Yes ^b | Yes | Yes ^b | Yes | Yes |
| Event no. | Yes | Yes ^c | Yes | Yes ^c | Yes | Yes |
| Risk table | No | No | No | Yes ^d | No | Yes ^e |
| Events annualised | No | Yes | No | Yes | No | No |
| Stepped graph | Yes | No | Yes | Yes | Yes | Yes |
| Median survival | No | No | No | No | No | No |

a Jungmann *et al.*¹¹¹ presented KM plots for time to end of follow-up and of time of failure for those who failed, so, as these did not allow analysis of time to failure, the study was excluded.

b Minas *et al.*¹³⁶ reported patient numbers only for the whole cohort, and for previous and no previous intervention subgroups. Nawaz *et al.*⁸⁰ reported patient numbers for whole cohort, for previous and no previous intervention subgroups, and for site of intervention subgroups, but not for grade of preoperative degenerative change subgroups.

c Minas *et al.*¹³⁶ and Nawaz *et al.*⁸⁰ event numbers were provided only for the whole cohort.

d A risk table was available only for the whole cohort.

e The risk table for the MF group was anomalous.

The definition of treatment failure varied between studies; failure either consisted exclusively of surgical re-intervention or a mixture of surgical re-intervention and a poor functional or pain score relative to pre treatment. *Table 7* summarises the treatment failure definitions used in the relevant studies.

TABLE 7 Failure definitions used in relevant studies

| Study | Definition of failure |
|--|---|
| Nawaz 2014 ⁸⁰ | (1) Graft delamination proven either by MRI or by arthroscopy; (2) a new surgical intervention, including arthroplasty, high tibial osteotomy or another revision procedure (graft hypertrophy was not counted as a failure); (3) a VAS pain score within ≤ 2 points of the preoperative score; or (4) a Stanmore functional score that was the same or worse than the preoperative score |
| Minas 2014 ¹³⁶ | (1) Graft failure with revision using partial knee arthroplasty or TKA; (2) graft failure with revision cartilage repair; and (3) graft survival but development of new defects elsewhere in the same knee necessitating additional surgery (progression of disease) |
| Moseley 2010 ¹¹⁵ | Needed an operation after ACI that necessitated removal of the graft, confirmed a loss of defect fill or violated the subchondral bone (e.g. abrasion chondroplasty, MF, drilling, unicompartmental knee replacement, TKR) |
| Vanlauwe 2011 ⁴³ | Re-intervention affecting more than 20% of the index lesion. TTF was the time between the end of the surgical procedure and 'the date of failure or re-intervention' |
| Knutsen 2007 ⁶⁷ | Failure if the patient needed a reoperation because of symptoms due to a lack of healing of the treated defect. The need for shaving or trimming of a lesion was not defined as a failure |
| Niemeyer 2014 ¹¹⁷ | Re-intervention surgeries |
| Jungmann 2012 ¹¹¹ | Time to event was defined as time to revision surgery |
| ACTIVE trial protocol www.active-trial.org.uk ³⁵ | Cessation of treatment benefit in which: 'two of the following three conditions below are satisfied: (a) Overall knee status judged by the assessor as not improved from preoperative condition (cessation of benefit form), (b) No gain in independently assessed Lysholm knee score compared with preoperative score, (c) No gain in patient's self-assessed Lysholm knee score compared with preoperative score'. Within element (a) re-intervention/additional procedures could be judged to be treatment failure |

TKR, total knee replacement; TTF, time to treatment failure; VAS, visual analogue scale.

Results of autologous chondrocyte implantation trials

Four small RCTs, each with two arms were identified. Two (Knutsen 2007,⁶⁷ $N = 80$ patients; Vanlauwe 2011,⁴³ $N = 112$ patients) compared ACI with MF. The RCT of Bentley *et al.* 2012⁷⁸ ($N = 100$ with about 10 years of follow-up) compared ACI with mosaicplasty – we are interested only in the ACI arm and these patients were included in the larger study of Nawaz *et al.*⁸⁰ and so are not considered further. The RCT of Gudas *et al.*¹²⁸ compared MF with mosaicplasty but did not satisfy inclusion criteria.

Figure 1 summarises the reconstructed time-to-failure KM plots for the Knutsen *et al.*⁶⁷ and Vanlauwe *et al.*⁴³ RCTs, which extended to 5 and 6 years of follow-up, respectively.

In these small studies the observed data are associated with considerable uncertainty. Extrapolation of parametric models was associated with large uncertainty beyond the observed data. Figure 2 illustrates this for Weibull fits extrapolated to 50 years (other models are presented in Appendix 9). The short follow-up and small size limits their usefulness for modelling failure rates beyond 5 years. Additional (non-RCT) studies of larger size and longer follow-up were sought.

Results of autologous chondrocyte implantation observational studies

Four single-arm ACI studies with KM plots were included. Niemeyer *et al.*¹¹⁷ reported event times for 70 German patients with follow-up to 5 years; Minas *et al.*¹³⁶ and Moseley *et al.*¹¹⁵ reported time to failure for 210 and 72 US patients, respectively, with follow-up extending up to or beyond 10 years; and Nawaz *et al.*⁸⁵ reported annualised time to failure for 827 UK patients, with follow-up to about 10 years. ACI patients from Bentley *et al.*⁷⁸ and from Biant *et al.*⁷⁹ were included in the study of Nawaz *et al.*⁸⁰ and so survival data from these are not considered separately. For the whole cohort, Nawaz *et al.*⁸⁰ reported both

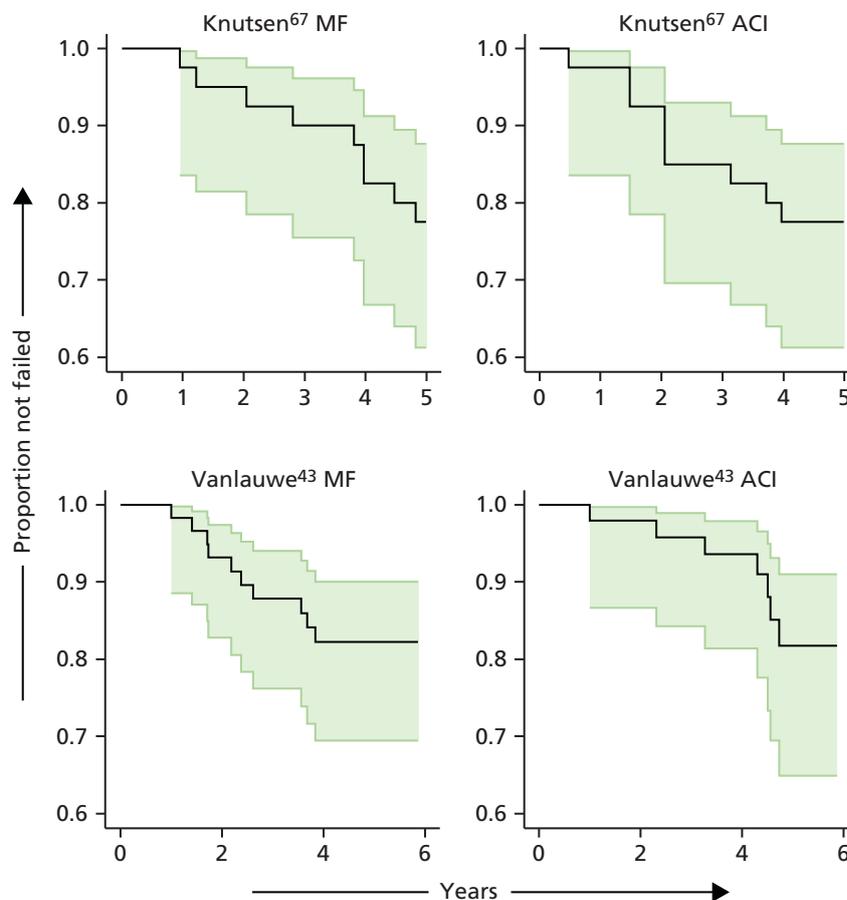


FIGURE 1 Reconstructed KM plots (95% CI) of time to failure in two RCTs.

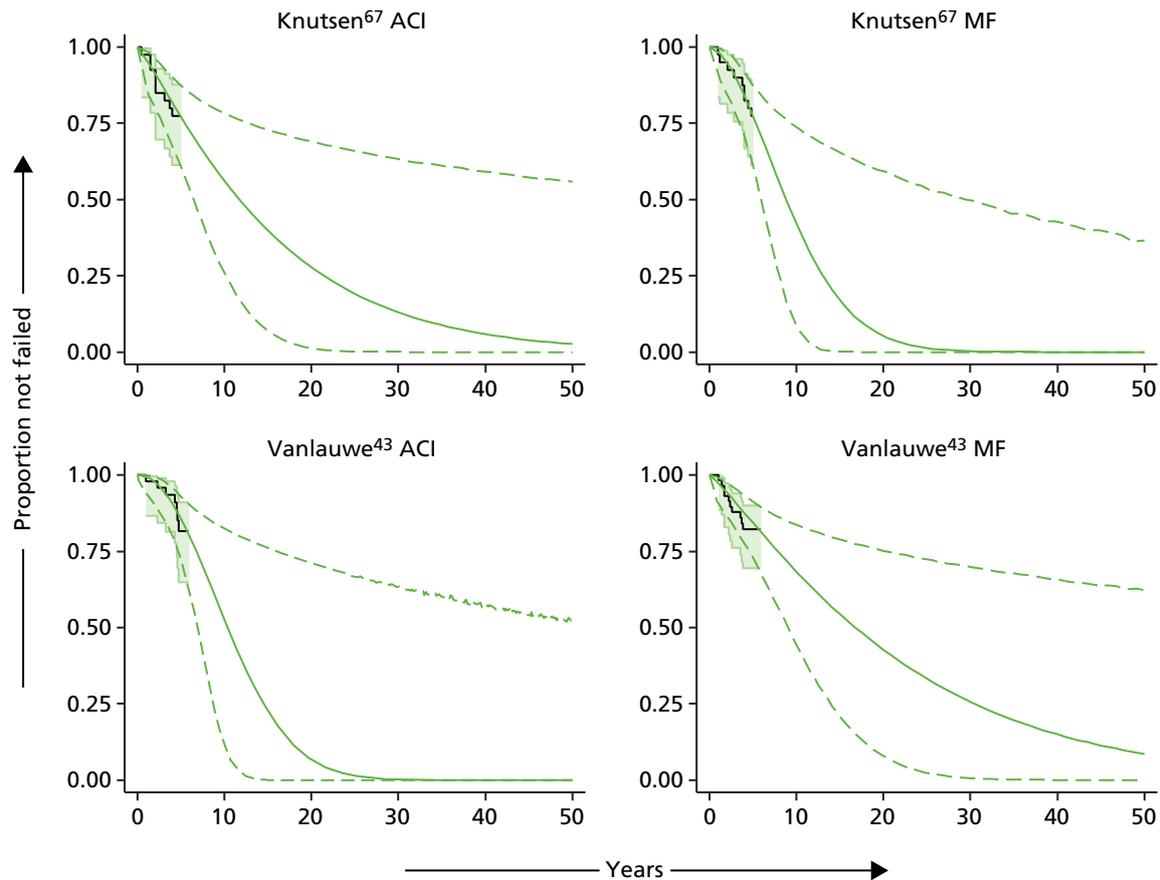


FIGURE 2 Extrapolated Weibull distributions fit to reconstructed IPD from the RCT studies. Note: there is some doubt about the reliability of the Vanlauwe *et al.*⁴³ MF arm because of an anomaly in the published risk table, which coincided with a flattening of the KM plot at 3 years. Saris *et al.* 2009⁶⁹ reported 3-year MF data for this trial.

annual event and censoring numbers for each year so that the IPD could be reconstructed without resort to the Guyot *et al.*¹³¹ algorithm. Jungmann *et al.*¹¹¹ presented time to re-intervention for 413 of 500 patients (selected follow-up 2–11.8 years) with analysis truncated at 5 years; this KM was not comparable with those in other studies and IPD were not reconstructed.

Patient characteristics in the ACI studies are summarised in *Table 8*. Typically, lesions were full thickness, with study mean size ranging from 2.7 to 8.4 cm² in patients with mean age 30–40 years, most of whom had experienced previous interventions. Symptom duration prior to intervention varied between studies.

Figure 3a shows the reconstructed KM failure plots for the four single-arm and two RCT ACI studies that provide relevant data to at least 5 years. It should be appreciated that definitions of failure were not identical between different studies, and the mix of patients who had, or had not, experienced previous intervention also differed between studies. Because of study size, the uncertainty in the Nawaz *et al.*⁸⁰ data is less than that in the other studies. Up to about 6 years there is reasonable consonance for most studies; thereafter, the prognosis appears worse for Nawaz *et al.*⁸⁰ patients, but this appears partly due to flat portions of the KM curves for the other studies for which patients at risk have diminished considerably and uncertainty is at its maximum. These data indicate that the Nawaz *et al.* study⁸⁰ is unlikely to flatter failure rates after ACI, and that to about 6 years of follow-up the Nawaz *et al.* study⁸⁰ is reasonably consistent with other studies; beyond 6 years the KM analyses of the other studies are likely to be less reliable because of smaller study size. Combination of these different studies would be difficult to justify because of clinical heterogeneity.

TABLE 8 Baseline characteristics reported for patients in six ACI studies

| Item | Bentley 2012 ⁷⁸ | Biant 2014 ⁷⁹ | Knutsen 2007 ⁶⁷ | Minas 2014 ¹³⁶ | Moseley 2010 ¹¹⁵ | Nawaz 2014 ⁸⁰ | Niemeyer 2014 ¹¹⁷ | Vanlauwe 2011 ⁴³ |
|---|----------------------------|--------------------------|----------------------------|---------------------------|-----------------------------|--------------------------|------------------------------|-----------------------------|
| <i>n</i> | 100 ^{a,b} | 104 ^b | 40 | 210 | 72 | 827 | 70 | 51 |
| FU ^c [range] | > 10 [10–12] | 10–12 | NR ^d | > 10 | 10 | 6.2 [2–12] ^e | 10.9, SD 1.1 | NR |
| Age ^c median [range] | 31.3 [16–49] | 30.2 [15–49] | 33.3 [NR] | 35.8 [8–57] | 37.0, SD 9.27 | 34 [14–56] ^e | 33.3, SD 10.2 | 33.9, SD 8.5 |
| Male (%) | 58 ^a | 52.9 | 60 | 53.8 | 61 | 59.6 | 35.7 | 61 |
| Defect size [range] (cm ²) ^e | 4 [1–10.5] | 4.8 [1.2–25] | 5.1 [NR] | 8.4, SD 5.5 | 5.2 [0.4–23.5] | 4.09 [0.64–20.7] | 6.5, SD 4.0 | 2.7 [1–5] |
| Previous (%) | 94 | 70 ^f | 93 | 42 | 74 | 34 | 62.8 | 88 |
| Mean no.: previous intervention(s) [range] | 1.5 | 1.3 [0–5] | 1.6 | NR | NR | NR | NR | NR |
| Weight [range] (kg) or BMI, kg/m ² | NR | NR | 81 [NR] | 26.7, SD 4.6 | 27.2 | NR | NR | 78.3, SD 13.9 |
| Symptom duration ^c [range] | 7.2 ^e [0.75–20] | 7.8 ^e | 3 ^g | NR | NR ^h | NR | Several | 1.97 ^g [0–18] |
| Defect site ⁱ | | | | NR | | | | |
| MF | 53.0 | 44.0 | 89.0 | | 72 ^j | 51 | 41.1 | 100 |
| LF | 18.0 | 16.0 | 11.0 | | 18 | 13 | 18.6 | 0 |
| Pa | 25.0 | 35.0 | 0 | | NR | 24.0 | 20 | 0 |
| Tr | 3.0 | 5.0 | 0 | | 10 | 6.0 | 2.9 | 0 |
| Mult | 0 | 0 | 0 | | NR | 6.0 | 17.1 | 0 |

FU, follow-up; LF, lateral femoral; MF, medial femoral; Mult, multiple; no., number; NR, not reported; Pa, patellar; SD, standard deviation; Tr, trochlea.

a ACI plus mosaicplasty groups.

b Consecutive patients.

c Years.

d Results reported for 5-year follow-up.

e Mean.

f Excludes debridement.

g Median.

h Sixty-five per cent had acute onset.

i Per cent.

j Seventeen per cent of patients in Moseley *et al.*¹¹⁵ had multiple sites.

Note

Six of 57 patients in Vanlauwe *et al.*⁴³ did not get treated.

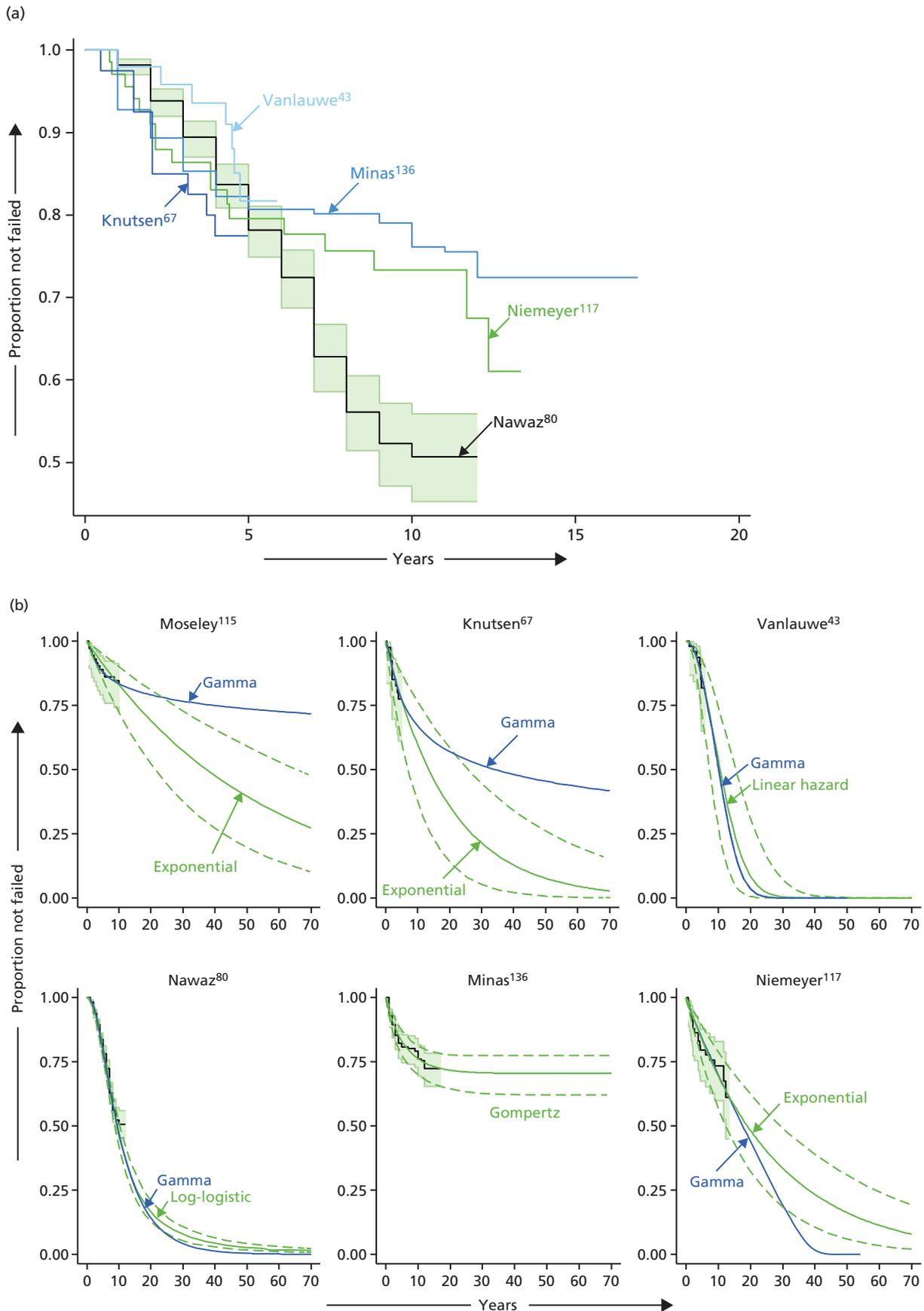


FIGURE 3 (a) Reconstructed KM plots for ACI studies; and (b) best parametric fit (95% CIs) and gamma models for six ACI arms.

Figure 3b shows parametric models for these studies extrapolated to 70 years. For clarity Moseley¹¹⁵ is omitted. A gamma model could not be computed for Minas *et al.*¹³⁶ Gamma fits illustrate differences seen between studies when applying the same single distribution to all; also shown are best fits for each study. Judged according to information criteria, various parametric models provided best fits: Knutsen⁶⁷ exponential, Minas¹³⁶ Gompertz, Moseley¹¹⁵ exponential, Nawaz⁸⁰ log-logistic, Niemeyer¹¹⁷ exponential and Vanlauwe⁴³ linear hazard. For Nawaz⁸⁰ and Vanlauwe⁴³ the best fits differed very little from the gamma model. The best fit models for Knutsen, Nawaz,⁸⁰ Niemeyer¹¹⁷ and Vanlauwe⁴³ studies predict that more than half of ACI interventions fail within about 30 years, whereas the Gompertz model-based Minas study¹³⁶ indicates about half or more ACIs remain without failure up to 70 years, and the gamma model for the Moseley study¹¹⁵ predicts about 25% remain without failure to 70 years.

Potential reasons for differences in KM plots and best fit models between these studies are manifold; they include uncertainty in the observed data resulting from small numbers of participants and, in some studies, short-term follow-up, different reliability of IPD reconstructions and differences in study populations, particularly with regard to experience of previous intervention(s), the degree of degenerative change, and location and size of lesion.

Post-failure treatments

Biant *et al.*⁷⁹ reported the revision surgeries following ACI failure as 44.4% total knee replacement (TKR) or unicompartmental knee replacement or patellofemoral joint replacement or medial and patellofemoral knee replacement; 25.9% ACI; 18.5% high tibial osteotomy; and 11.1% arthrodesis or chondroplasty. In the Minas *et al.* study,¹³⁶ 19 out of 53 patients with failed grafts went on to knee arthroplasty within the follow-up period, 27 out of 53 patients had revision cartilage repair procedures and 7 out of 53 patients refused further treatment after failure.

Studies with patient subgroup analyses

Jungmann *et al.*¹¹¹ and Bentley *et al.*⁷⁸ reported data (but not KM plots) comparing failure rates between subgroups of patients. Jungmann provided evidence that increased revision was associated with previous intervention, previous bone marrow stimulation, female gender and ACP-P relative to other ACI types. Bentley provided 5-year revision rates by subgroup; only older age appeared associated with increased probability of revision (note that these patients were included in the Nawaz study⁸⁰).

The Minas and Nawaz studies^{80,136} both presented KM plots for subgroups of patients, but neither reported event numbers by subgroup, and patient numbers were available only for some subgroups of patients. Nawaz *et al.*⁸⁰ provided Cox regression HRs for several subgroups of patients. Because of its size, length of follow-up, the use of multiple surgeons and inclusion of UK patients, the Nawaz study⁸⁰ was judged to be the most relevant ACI study for the current decision problem. Therefore, the focus in this section is on the Nawaz study,⁸⁰ and the results from the Minas study¹³⁶ are presented for comparison.

Nawaz *et al.* 2014 study of UK patients

The most useful study is by Nawaz *et al.*⁸⁰ For the whole Nawaz cohort ($n = 827$)⁸⁰ a log-logistic distribution provided the best-fitting parametric model. Figure 4 shows the reconstructed KM plot together with the log-logistic model extrapolated to 50 years; the model predicts that after about 30 years approximately 90% of patients would have failed. The partition of failures according to how it was defined (see Table 7) was not reported (e.g. the proportion of failures receiving previous intervention at the time of failure is unknown).

Patient subgroups examined in the Nawaz study

Nawaz *et al.*⁸⁰ presented KM plots for subgroups of patients categorised according to (1) receipt of a previous intervention; (2) site of intervention; (3) grade of degenerative change; and (4) type of intervention received (MACI or ACI). The authors used univariate and multivariate Cox regression to investigate if these – and also age and size of defect – were influential for failure. The most influential patient covariate was previous intervention ($p < 0.001$; multivariate HR vs. no previous intervention: 4.72, 95% CI 3.5 – 6.4). Grade of degenerative change ($p < 0.001$), site of intervention ($p = 0.036$ for best vs. worst site), and age at operation ($p < 0.001$) were also significantly influential, whereas type of intervention (ACI or MACI) and

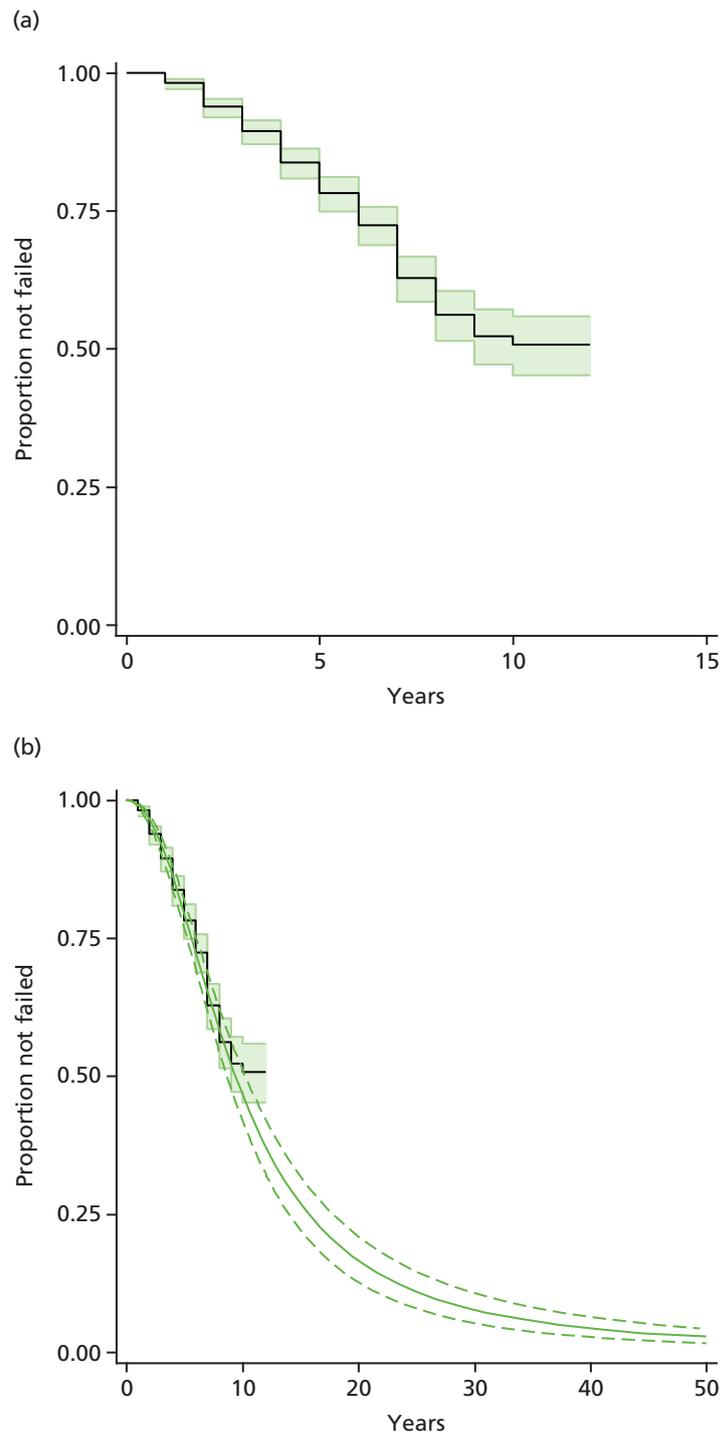


FIGURE 4 Reconstructed KM plot (a) and extrapolated log-logistic model; and (b) for the Nawaz *et al.*⁸⁰ whole cohort.

lesion size were not ($p = 0.860$ and $p = 1.00$, respectively). The authors did not report on a test of the proportional hazards assumption. The AG reconstructed the subgroup KM plots and used reconstructed IPD to investigate good parametric models for the data. Additionally, AG investigated the effect of adjusting parametric models using the multivariate HRs reported by Nawaz *et al.*⁸⁰

Previous and no previous intervention

According to information criteria, log-normal and gamma distributions provided good models for patients who had previous or no previous intervention (debridement was not included as a previous intervention). When the HR reported by Nawaz *et al.*⁸⁰ (previous vs. no previous intervention) was applied to either

log-normal or Weibull models, the resulting model was very similar to that fit to the previous subgroup IPD (Figure 5). These results indicate that there was likely to be little difference between the subgroups in the distribution of other covariates that were influential for failure.

Site of intervention

Nawaz *et al.*⁸⁰ published KM plots and multivariate Cox regression HRs comparing time to failure for five subgroups that differed according to intervention site (medial femoral, $n = 421$; lateral femoral, $n = 109$; patella, $n = 200$; trochlea, $n = 50$; multiple sites, $n = 47$). HRs vs. the lateral femoral condyle group as baseline reference were medial femoral condyle 1.806 (95% CI 1.036 to 3.149; $p = 0.037$), patella 1.323

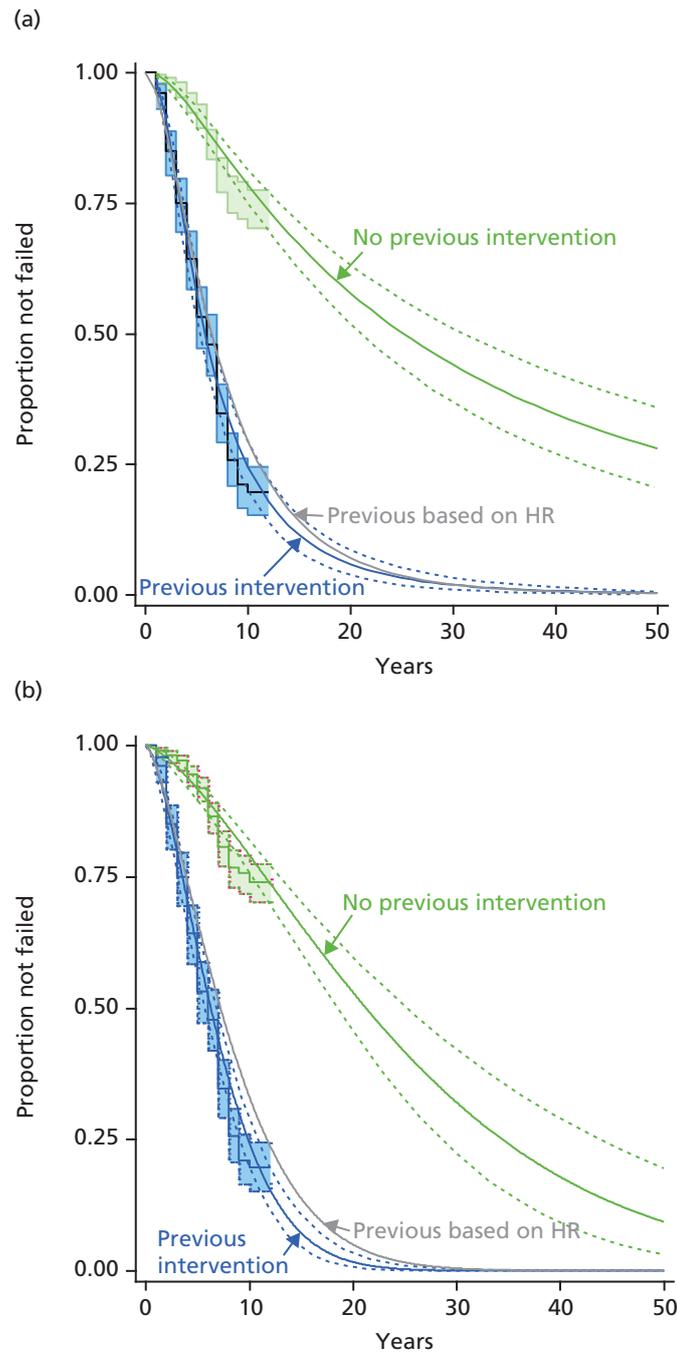


FIGURE 5 Reconstructed plots for Nawaz *et al.*⁸⁰ Previous and no previous intervention subgroups: (a) log-normal models; and (b) Weibull models. KM plots and extrapolated log-normal (left) and Weibull (right) parametric models. Dashed lines are 95% CIs.

(95% CI 0.745 to 2.351; $p = 0.339$), trochlea 1.409 (95% CI 0.625 to 3.174; $p = 0.0408$) and multisite 1.678 (95% CI 0.731 to 3.851; $p = 0.222$). Reconstructed KM plots were similar for all but the lateral condyle group, which exhibited the least failure. Log-normal distributions provided the best fit parametric models to reconstructed subgroup IPD. *Figure 6* shows reconstructed KM plots and HR-adjusted log-normal models. Applying the reported HRs diminished the apparent superiority of the lateral femoral condyle subgroup seen in the KM plots, and indicated that relative to other subgroups the lateral femoral population may possibly have been favourably free of detrimental covariates for failure (e.g. previous treatment and high-grade degenerative change). Similar results were obtained with Weibull models.

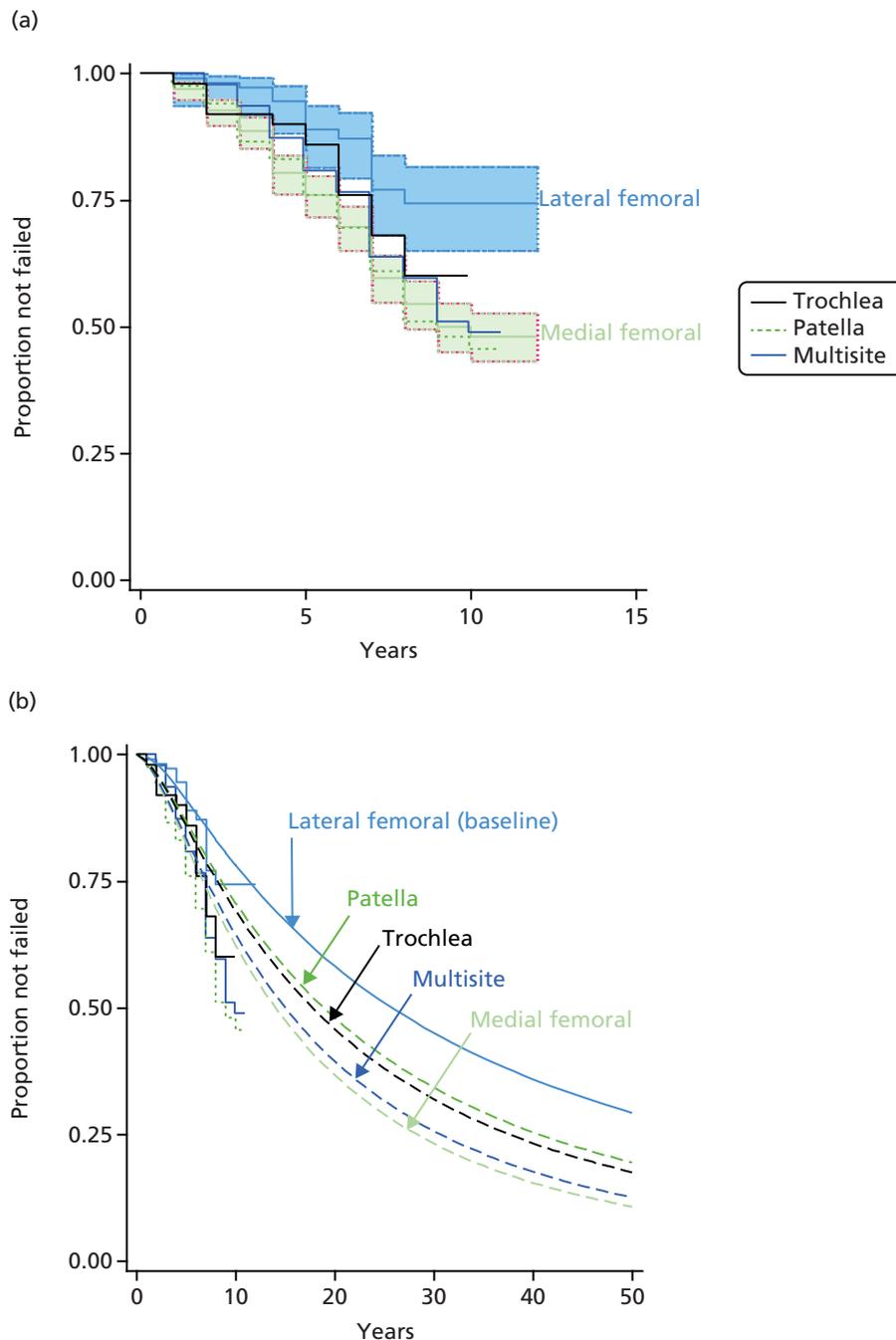


FIGURE 6 Reconstructed KM plots and log-normal models for time to failure according to site of intervention, Nawaz *et al.*⁸⁰ (a) KM according to intervention site; and (b) subgroups by intervention site log-normal models based on HRs. For clarity, KM 95% CIs are shown only for lateral and medial femoral condyle sites and have been omitted for the model fits.

Grade of preoperative degenerative change

Nawaz *et al.*⁸⁰ published KM plots and multivariate Cox regression HRs comparing time to failure for four subgroups categorised according to grade of degenerative change (Kellgren–Lawrence). HRs vs. the grade 0 subgroup as baseline reference were grade 1, 1.542 (95% CI 0.930 to 2.557; $p = 0.093$); grade 2, 1.869 (95% CI 1.381 to 2.529; $p < 0.001$); and grade 3, 1.985 (95% CI 1.092 to 3.610; $p = 0.025$). Patient numbers were not reported and parametric models were fit to digitised KM plots using non-linear regression. For different subgroup grades, log-normal and linearly increasing hazard models produced acceptable fit to digitised KM plots in *Appendix 9*. When the HRs reported by Nawaz *et al.*⁸⁰ were applied to either of these models, the resulting plots for different grades were more similar to each other than was apparent from KM plots or fits to KM plots (*Figure 7*). These results may indicate that some of the superiority of the grade 0

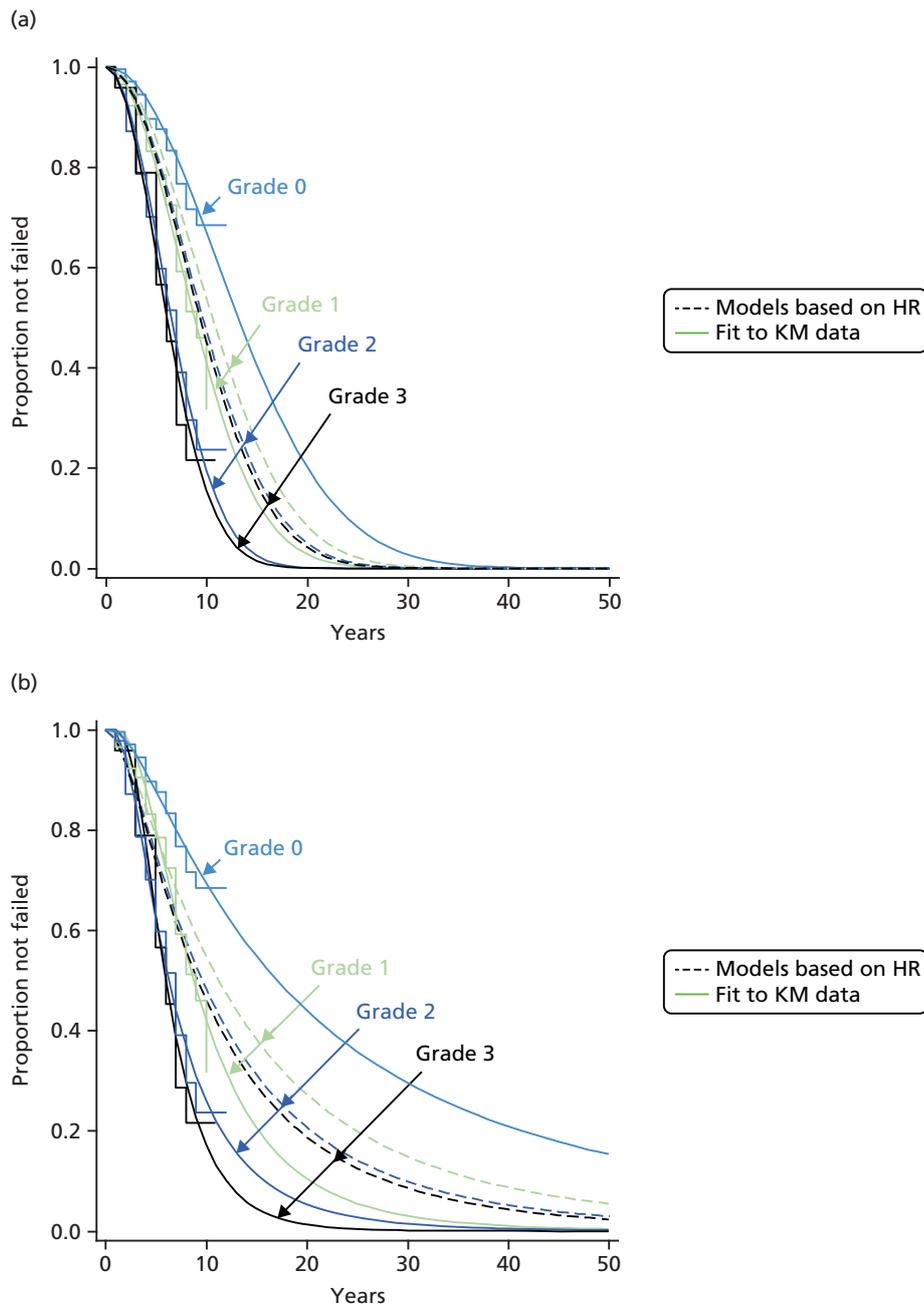


FIGURE 7 Time to failure by grade of OA. (a) Nawaz *et al.*⁸⁰ subgroups by grade of degenerative change linearly increasing hazard models; and (b) Nawaz *et al.*⁸⁰ subgroups by grade of degenerative change log-normal models.

subgroup apparent in the KM plots was possibly due to relative freedom of this group from covariates that tend to increase the probability of failure.

Minas et al. 2014 study

Minas *et al.*¹³⁶ performed several KM analyses for various subgroups of patients. Patient numbers were reported only for the comparison of previous intervention versus no previous intervention groups. Like in the Nawaz *et al.* study,⁸⁰ worse failure rates were found for patients who had experienced previous intervention. As was seen for the whole Minas cohort,¹³⁶ the subgroup failure rates flattened after about 6 years and extended to as far as 17 years with relatively few failure events. Thus failures were much less frequent in both Minas¹³⁶ subgroups than in the corresponding Nawaz⁸⁰ subgroups. No regression analysis was performed in Minas¹³⁶ and no HRs were reported. Gamma distributions provided good fits for both studies' subgroups. Reconstructed KM plots and gamma models fit to IPD for subgroups from both the Nawaz⁸⁰ and Minas¹³⁶ studies are shown in *Figure 8*.

Minas *et al.*¹³⁶ also provided plots for failure according to subgroups that experienced different types of previous intervention. Patient numbers were as follows: MF, $n = 13$; abrasion arthroplasty, $n = 30$; and drilling, $n = 46$. Failure was more frequent after MF than after other forms of marrow stimulation. Concurrent osteotomy resulted in fewer failures.

Studies of failure after microfracture (Layton 2015, Knutsen 2007, Saris 2009)

Vanlauwe *et al.*⁴³ (year 3 results in Saris *et al.*⁷⁰) and Knutsen *et al.*⁶⁷ provided MF failure data to 5 years. A large US study by Layton *et al.*¹¹⁴ examined records for 3498 US recipients of MF and reported the percentages of failures for patients followed to 1, 3 and 5 years. All patients were followed up to 3 years. The proportion followed to 5 years was not reported. Layton *et al.* stated 'Failure rates (TKR, MF or ACI) increased with increasing years of follow-up: 9% within 1 year, 18% within 3 years, and 32% within 5 years'. In the Knutsen,⁶⁷ Saris⁷⁰ and Layton¹¹⁴ studies, failure was defined as re-intervention. Only Layton *et al.*¹¹⁴ provided information on the type of re-intervention received, as follows: TKR accounted for most re-interventions – 56%, 62% and 66% of re-interventions at years 1, 3 and 5, respectively; MF and ACI

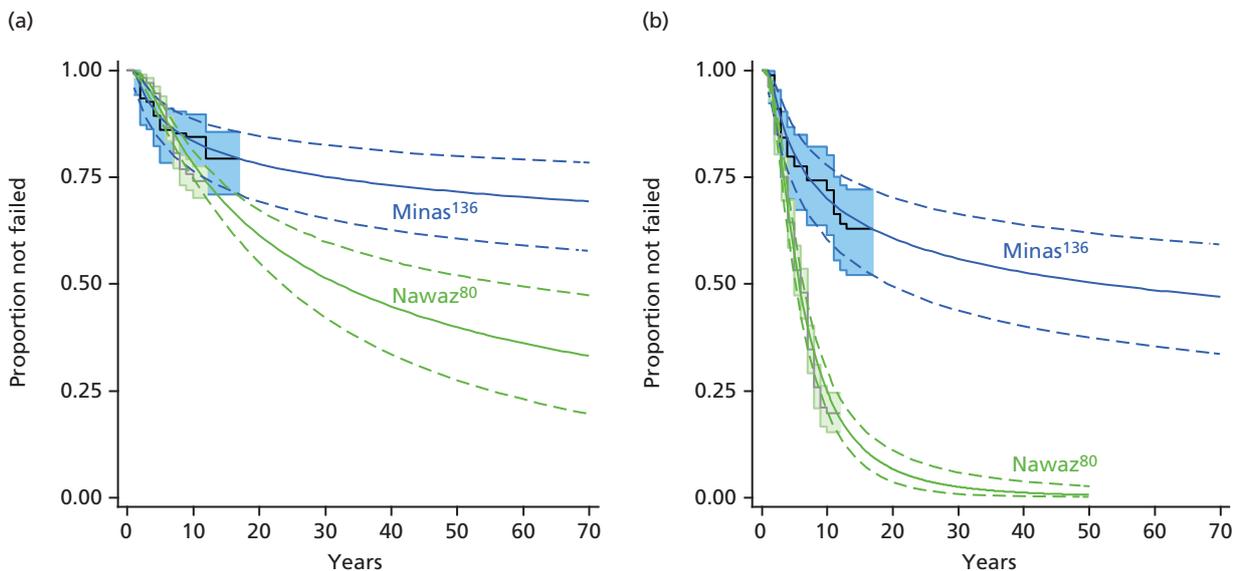


FIGURE 8 Effect of previous interventions on failure rates. (a) Nawaz *et al.*⁸⁰ and Minas *et al.*¹³⁶ no previous intervention populations. Reconstructed KM plots and gamma models; and (b) Nawaz *et al.*⁸⁰ and Minas *et al.*¹³⁶ previous intervention populations. Reconstructed KM plots and gamma models. Time to failure for previous intervention and no previous intervention patient subgroups (Minas¹³⁶ and Nawaz⁸⁰) showing reconstructed KM plots and gamma models of time to failure.

accounted for nearly all of the remaining re-interventions (very few re-interventions were OATS). The mean age of patients in the Layton study¹¹⁴ was 47 years [standard deviation (SD) 11.4 years], meaning that many would be of an age at which TKR would be considered, and there were equal numbers by gender. *Table 9* summarises the main characteristics of patients in the MF arms of the Knutsen⁶⁷ and Saris⁷⁰ studies.

As all of the patients in the Layton study¹¹⁴ were followed up for 3 years, it was possible to reconstruct IPD for an annualised KM plot (assuming failure took place at 1 and 3 years, and at 3 years all non-failed patients were censored). Under the assumption that those followed up for 5 years were representative of all of those who could have been followed (censoring those without failure at 5 years), the 5-year IPD were also estimated. The best fit models (*Figure 9*) for these were provided by Gompertz distributions and the second best by a gamma model (see *Appendix 9*).

The linearly increasing hazard model provided the best fit for the MF arm of the Knutsen *et al.*⁶⁷ study.

The published 5-year KM plot for the MF arm of the Vanlauwe *et al.* study⁴³ had anomalous risk table data and interpretation of the KM plot was problematic. The KM plot for MF has 10 steps and a total of 10 events were reported (one step for each event). Seven steps occur before 36 months: two of these very close together, at about 20 months, and three after 36 months. This does not tally with the data in their appendix 1, which depict five MF re-interventions occurring before 36 months and five after 36 months. For the ACI KM plot, two steps occur before 36 months and five steps after 36 months, and this corresponds to the data provided in appendix 1 of the Vanlauwe *et al.* study.⁴³ The risk table for the MF arm is anomalous in that the number at risk is reported as increasing at 36 months. It is unclear what the correct numbers should be at 24, 36 and 48 months for the MF risk table. Therefore, the Saris *et al.*⁷⁰ 3-year KM plot for this study was examined. The best fit was again provided by the linearly increasing hazard model. These and models for the Layton *et al.* study¹¹⁴ are summarised in *Figure 9*. The poorer performance in the Layton *et al.* study¹¹⁴ may be attributable to older mean age and or real-world performance of MF relative to that for patients who are carefully selected for a RCT.

TABLE 9 Patient characteristics in the MF arms of the Saris⁷⁰ and Knutsen⁶⁷ RCTs

| Item | Knutsen 2007 ⁶⁷ | Saris 2009 ⁷⁰ |
|--|----------------------------|--------------------------|
| No. | 40 | 61 |
| FU (years) | 5 | 5 |
| Mean (SD) age (years) | 31.1 (NR) | 33.9 (8.6) |
| Male (%) | NR | 67 |
| Mean (SD) defect size (cm ²) | 4.5 (NR) | 2.4 (1.2) |
| Mean (SD) weight (kg) | 82.1 (NR) | 80.6 (13.3) |
| Mean number previous operations | 1.4 | NR |
| Previous operation | 93% | 77% |
| Median symptom duration, years | 3 | 1.57 [range 0–18] |
| Site medial femoral | 89% | NR |
| Site lateral femoral | 11% | NR |

FU, follow-up; NR, not reported.

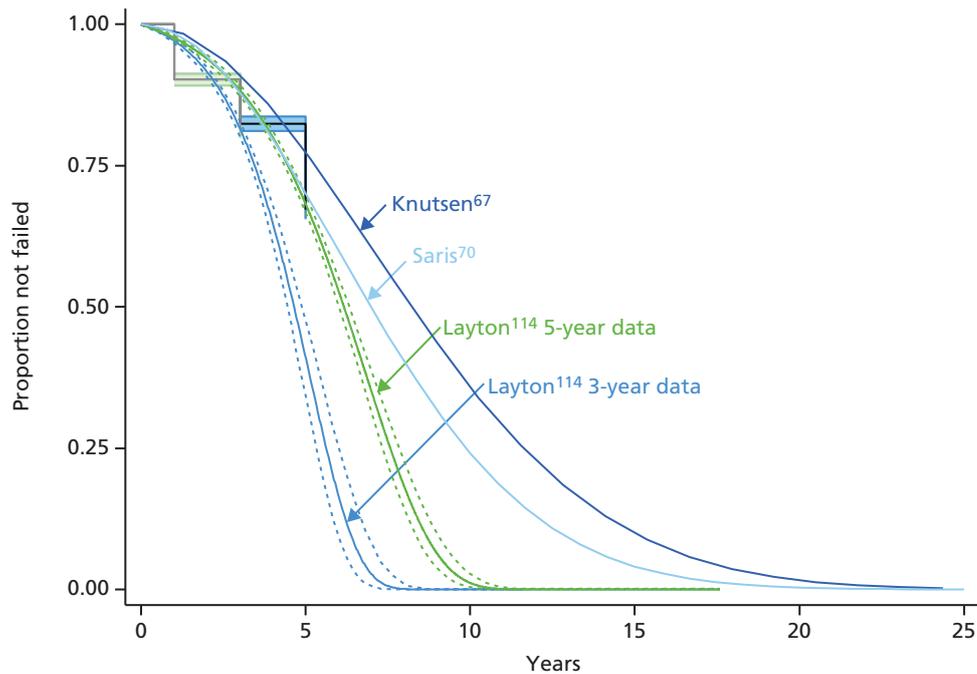


FIGURE 9 Best fit models to observed MF failure in three studies.

Comparison of failure after autologous chondrocyte implantation and microfracture

A comparison of long-term failure of MF and ACI is problematical in view of the paucity and heterogeneity of studies. The most reliable comparison may be between the largest UK extended follow-up study (Nawaz *et al.*⁸⁰) and the available MF data (described above), a caveat being that failure definitions differed between the Nawaz study⁸⁰ and the three available MF studies. When whole cohorts were compared, ACI appears to be superior to MF (Figure 10).

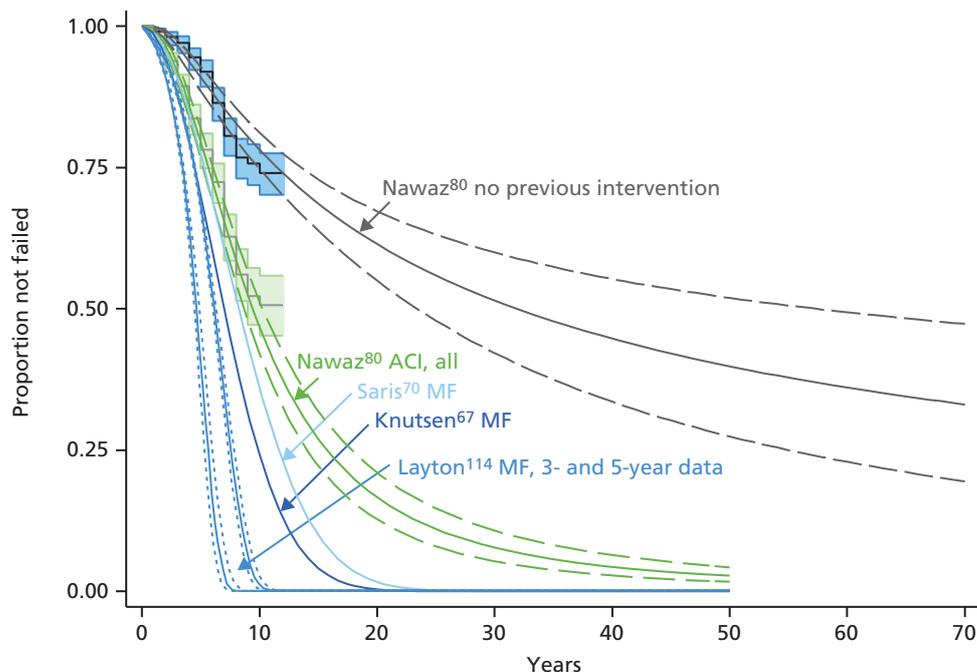


FIGURE 10 Modelled failure profiles following ACI or MF. Best fit models for Nawaz *et al.*⁸⁰ whole cohort, (all) log-log; Nawaz *et al.*⁸⁰ no previous subgroup gamma; Knutsen *et al.*⁶⁷ and Saris *et al.*⁷⁰ MF arms linearly increasing hazard; and Layton *et al.*¹¹⁴ Gompertz (for clarity not all 95% CIs are shown).

No subgroup data was available from the MF studies. Vanlauwe *et al.*⁴³ did not provide KM plots for subgroups but reported the failure numbers according to whether or not previous intervention had been experienced (Table 10). Numbers of patients at risk and numbers of events were small and the time of events in compared groups was not provided so that firm conclusions are impossible; however, these data are suggestive of little effect of previous intervention on risk of failure after either ACI or MF. Salzmann *et al.*¹³⁷ followed 454 recipients of MF and compared patient characteristics between those patients who required re-intervention during follow-up with those who did not require re-intervention. The former patients, on average, had received more pre-MF interventions (1.9 ± 2.1 previous interventions) than the latter (1.2 ± 2.1 previous interventions) but the spread in number of pre-interventions was great in both cases. Unfortunately, no KM time-to-event analyses were reported for the 'no-previous intervention' and 'previous intervention' subgroups.

In the absence of subgroup KM data for MF, the worst-performing subgroups investigated by Nawaz *et al.*⁸⁰ were compared with the three MF studies. Log-normal models based on the multivariate HRs reported by Nawaz *et al.*⁸⁰ were used for the comparison (Figure 11).

Except for the previous intervention ACI subgroup, the ACI subgroups clearly exhibited less failure than the MF cohorts. Lack of data does not allow a comparison with previously treated MF patients. It should be emphasised that uncertainty in these comparisons is substantial, especially with regard to the Knutsen⁶⁷

TABLE 10 Failure of ACI and MF according to previous intervention (data from Vanlauwe *et al.* 2011⁴³)

| Type of intervention | ACI failures/group (risk of failure) | MF failures/group (risk of failure) |
|-----------------------------|--------------------------------------|-------------------------------------|
| Previous knee surgery | 6/50 (0.120) | 8/47 (0.170) |
| No previous knee surgeries | 1/7 (0.143) | 2/14 (0.143) |
| 1 previous knee surgery | 3/29 (0.103) | 4/34 (0.118) |
| ≥ 2 previous knee surgeries | 3/21 (0.143) | 4/13 (0.308) |

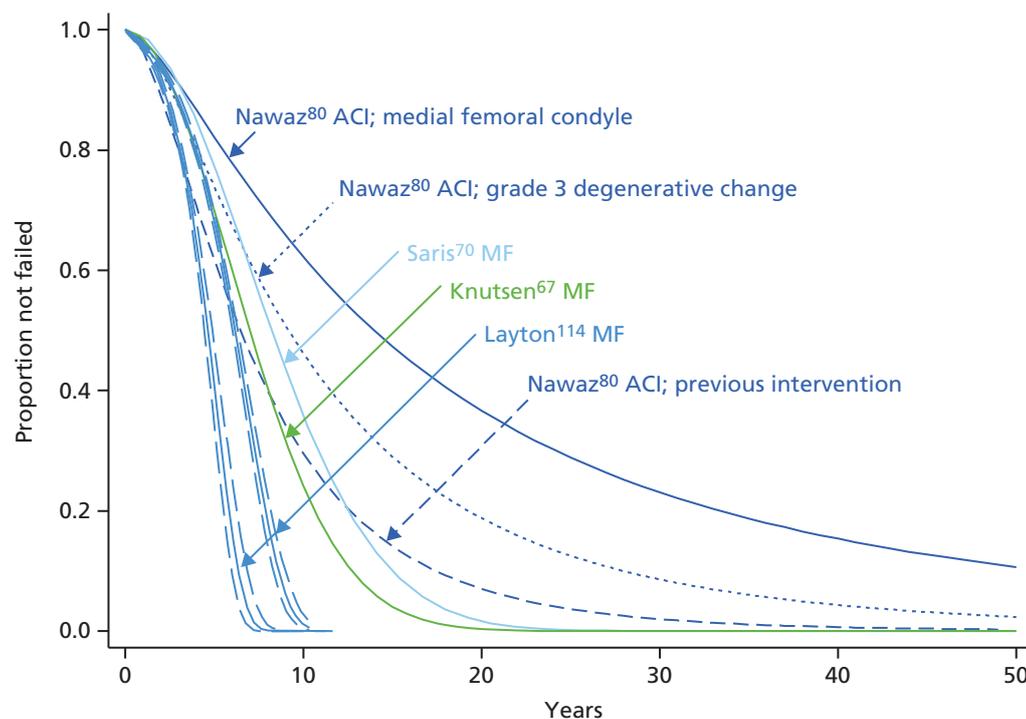


FIGURE 11 Modelled failure in MF studies compared with worst-performing subgroups from Nawaz *et al.*⁸⁰ ACI study.

and Saris⁷⁰ MF arms. Analysis of the MF arm of the RCT of Gudas *et al.*¹²⁸ – excluded on the basis of its small size – showed that the best fit for the reconstructed IPD was a log-normal model, which predicted poorer survival than the models for MF for the Layton,¹¹⁴ Knutsen⁶⁷ and Saris⁷⁰ studies.

Summary of longer-term time-to-event evidence of treatment failure after autologous chondrocyte implantation and microfracture

- More long-term evidence was available for ACI than for MF.
- Treatment failure definitions differed between studies, with varying, and sometimes unclear, relative contributions to overall failure from re-intervention and from inadequate pain/function scores.
- Study data were generally still too short term. Only one published study allowed an estimate of observed median time to failure.
- Most participants in most study populations had experienced intervention(s) prior to enrolment. When evidence was reported it appears many types of pre-intervention had been tried.
- Two ACI studies with KM analyses extending to at least 10 years reported that treatment failure was far more frequent in patients who had experienced prior intervention(s); one of these studies documented greater failure rates after MF than after other marrow stimulation (but patient numbers were small).
- There was no clear time-to-event evidence that prior intervention influenced failure after MF; other available evidence was meagre.
- Immaturity of failure data necessitated parametric modelling beyond observed data so as to predict lifetime failure.
- According to information criteria and visual goodness of fit, the best fits of long-term failure after ACI were usually characterised by models that, when extrapolated beyond the observed data, indicated gradually decreasing hazard (probability of failure decreasing with time).
- Conversely, good fits to limited data available for MF were characterised by models that indicated linearly increasing hazard (probability of failure increasing with time).
- A single large US study of MF in patients with a mean age of 47 years indicated that, in this population, TKR was the most frequent intervention after failure of MF.

Pooling time-to-failure studies

The second submission from SoBi used parametric models based on pooled data from ACI studies to derive time to failure for ACI. SoBi did not pool MF studies. A commentary on the SoBi submission follows in *Chapter 4*.

The ACI studies pooled by SoBi encompassed studies using different definitions of failure and recruiting different proportions of previously treated and previously untreated patients. More judicious pooling can be undertaken, in which there is less heterogeneity among pooled studies. Therefore, as a supplement to the analysis of single studies described above, the AG have briefly explored pooling of studies for ACI and for MF.

Autologous chondrocyte implantation studies

In the ACI arms of the studies by Moseley *et al.*,¹¹⁵ Vanlauwe *et al.*,⁴³ Knutsen *et al.*⁶⁷ and Niemeyer *et al.*,¹¹⁷ failure was defined as re-intervention and each study included more than 60% of patients who had experienced previous intervention (range 63–90%). A log-normal model provided the best fit for these pooled studies (*Figure 12*).

Microfracture studies

In the three studies providing MF data, failure was defined as re-intervention and two predominantly recruited patients who had experienced previous interventions. This was not reported by Layton *et al.*¹¹⁴ When the three studies (Layton,¹¹⁴ Vanlauwe⁴³ and Knutsen⁶⁷) were pooled the resulting KM plot and best fit model (Gompertz) were dominated by the large Layton study¹¹⁴ (see *Figure 12*). Compared with the pooled ACI studies, failure was more frequent in the MF studies.

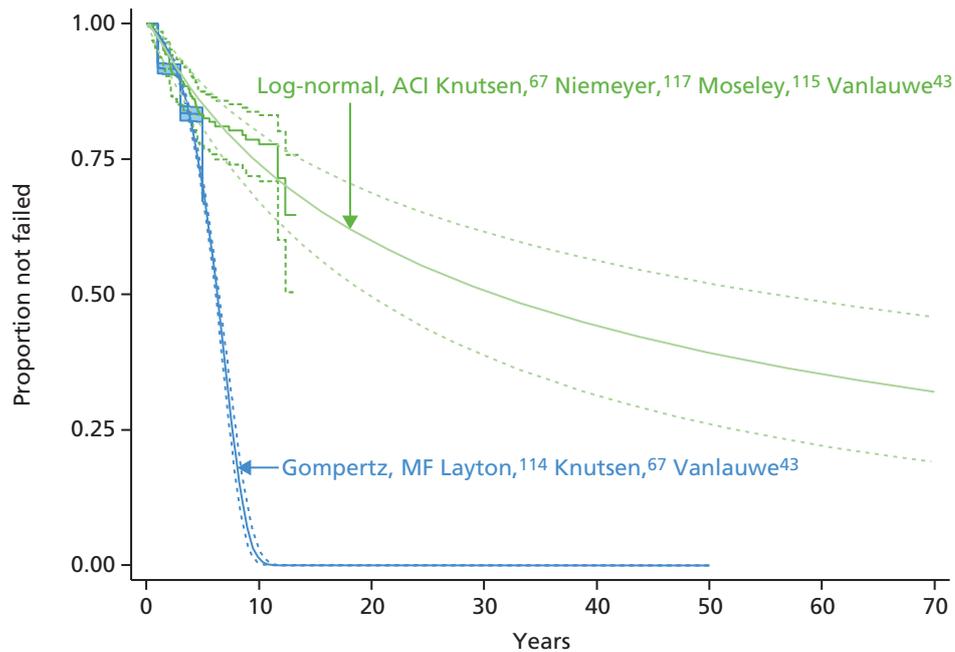


FIGURE 12 Time to failure after pooling four ACI studies and three MF studies.

The pooled MF studies were dominated by the Layton study.¹¹⁴ Pooled MF studies, excluding Layton,¹¹⁴ Knutsen⁶⁷ and Saris,⁷⁰ again indicated less failure for ACI patients than for MF patients (*Figure 13*).

When the MF arms of Knutsen⁶⁷ and Vanlauwe⁴³ (5-year MF data of the TIG/ACT study⁴³) were pooled, it was difficult to determine the best fit model using information criteria (*Table 11*). Only the gamma model of MF failure was superior to ACI (*Figure 14*). It should be noted that anomalies in the published Vanlauwe⁴³ MF arm required speculative interpolation of risk table data prior to pooling.

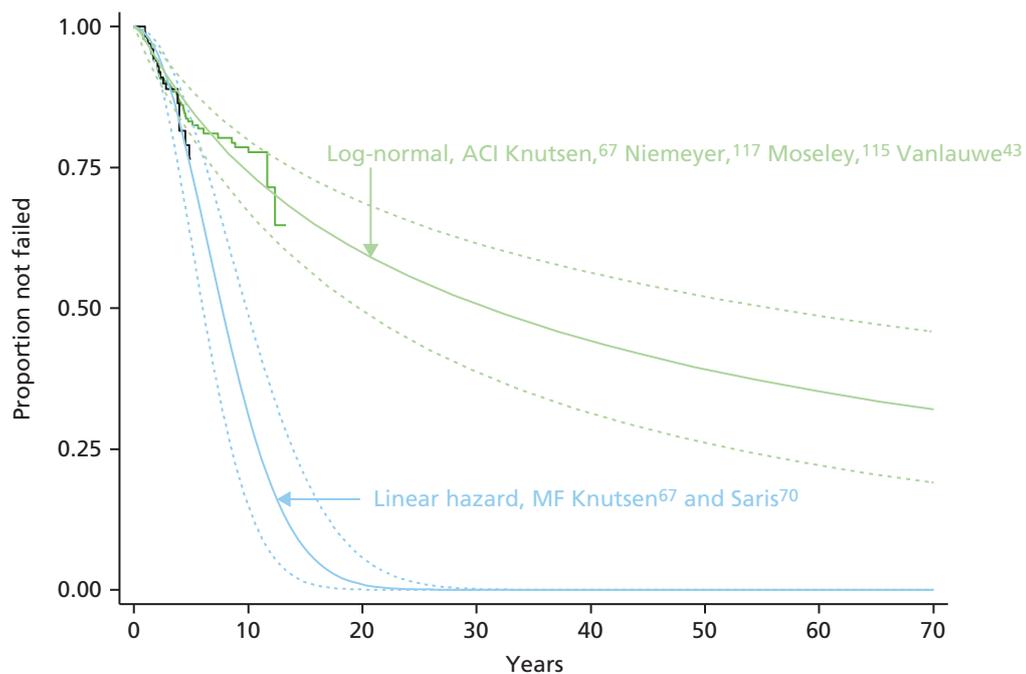
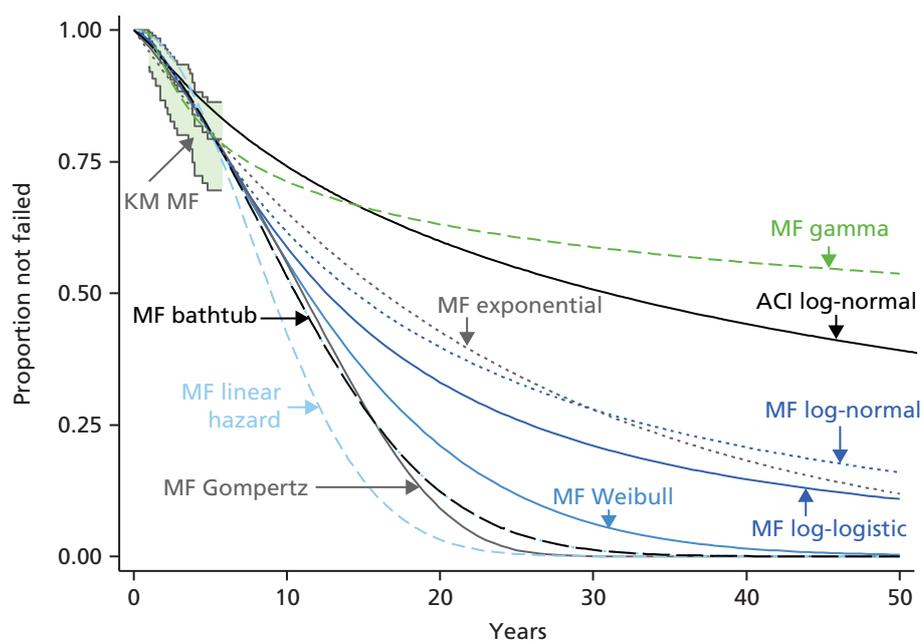


FIGURE 13 Time to failure after pooling four ACI and two MF studies (Knutsen⁶⁷ and Saris⁷⁰).

TABLE 11 Information criteria for models to pooled MF data from Knutsen⁶⁷ and Vanlauwe⁴³

| Model | Obs. | ll(model) | df | AIC | BIC | AIC rank | BIC rank |
|------------------------------|------|-----------|----|----------|----------|----------|----------|
| Gamma | 101 | -58.4314 | 3 | 122.8628 | 130.7081 | 1 | 5 |
| Exponential | 101 | -61.9545 | 1 | 125.9089 | 128.524 | 5 | 1 |
| Weibull | 101 | -60.7536 | 2 | 125.5071 | 130.7374 | 4 | 6 |
| Gompertz | 101 | -61.6625 | 2 | 127.325 | 132.5553 | 8 | 8 |
| Log-normal | 101 | -59.7245 | 2 | 123.449 | 128.6793 | 2 | 2 |
| Log-logistic | 101 | -60.5069 | 2 | 125.0139 | 130.2441 | 3 | 4 |
| Linear hazard, one parameter | 101 | -62.1378 | 1 | 126.2756 | 128.8907 | 6 | 3 |
| Bathtub | 101 | -61.4594 | 3 | 128.9189 | 136.7642 | 9 | 9 |
| Linear hazard, two parameter | 101 | -61.4594 | 2 | 126.9189 | 132.1491 | 7 | 7 |

df, degrees of freedom; Obs., observations.

**FIGURE 14** Time to failure after pooling four ACI and two MF studies (Knutsen⁶⁷ and Vanlauwe⁴³).

Chapter 4 Systematic review of existing economic studies for autologous chondrocyte implantation

Introduction

The objective of this chapter was to conduct a systematic review of existing economic evaluations (including any model-based economic evaluations) of the use of ACI, MF and mosaicplasty for repairing symptomatic articular cartilage defects of the knee. We searched the literature since the last HTA review³ for economic evaluations, including any existing models, to help inform our economic modelling.

Methods

The systematic search used MEDLINE Ovid (2004 to 6 July 2014), EMBASE Ovid (2004 to 6 July 2014), NHS Economic Evaluation Database (NHS EED; issue 2 of 4, April 2014) and the Web of Science Core Collection (2004 to 6 July 2014). Weekly auto-alerts were set up in Ovid MEDLINE and EMBASE for any new studies that were added to the database subsequent to July 2014. The search terms included economic and QoL terms cross-referenced with chondrocyte implantation terms. The search was limited to studies published since the searches were done for the last HTA review,³ that is from the year 2004. The search was also limited to studies published in the English language and to humans. Details of the search strategies are provided in *Appendix 10*.

Two reviewers independently reviewed titles and abstracts to identify potentially relevant papers. Consensus was achieved by discussion, but, when consensus was not agreed, a third reviewer reviewed the abstracts to reach agreement. Abstracts were considered to be relevant to this review if they were a full economic analysis (including any economic models) on the use of ACI, MF and mosaicplasty for repairing symptomatic articular cartilage defects of the knee. Abstracts that provide useful information for the economic model (such as costs, utilities and transition probabilities) were retained but not included in this review.

We obtained the full-text articles of potentially relevant abstracts. The reference lists of retrieved articles were checked for potentially relevant papers that met the inclusion criteria. A data extraction form was developed to capture the main characteristics and economic factors. We critically appraised full economic evaluations against the framework for quality assessment of economic evaluation studies developed by the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) group.¹³⁸ If the studies contained an economic model, they were further assessed against a framework for the quality assessment of decision-analytic modelling adapted from Philips *et al.*¹³⁹

Results

The searches identified 272 potentially relevant citations published since 2004. After reviewing the abstracts, four studies remained, including the HTA review by Clar *et al.* 2005³ (Derrett 2005,¹⁴⁰ Gerlier 2010,¹⁴¹ Samuelson 2012¹⁴²). A further two articles were identified from the clinical effectiveness searches (MSAC 2010,³² Koerber 2013¹⁴³). In total, six articles were retained for data extraction. *Figure 15* shows the abstracts identified and number of studies included.

Of the six publications, two have been summarised (see below). The Clar *et al.* study³ is based on previous work by some of the authors of this report. The MSAC report published in 2010³² compared only the costs, as MSAC assumed that the clinical effectiveness for the different interventions was identical.

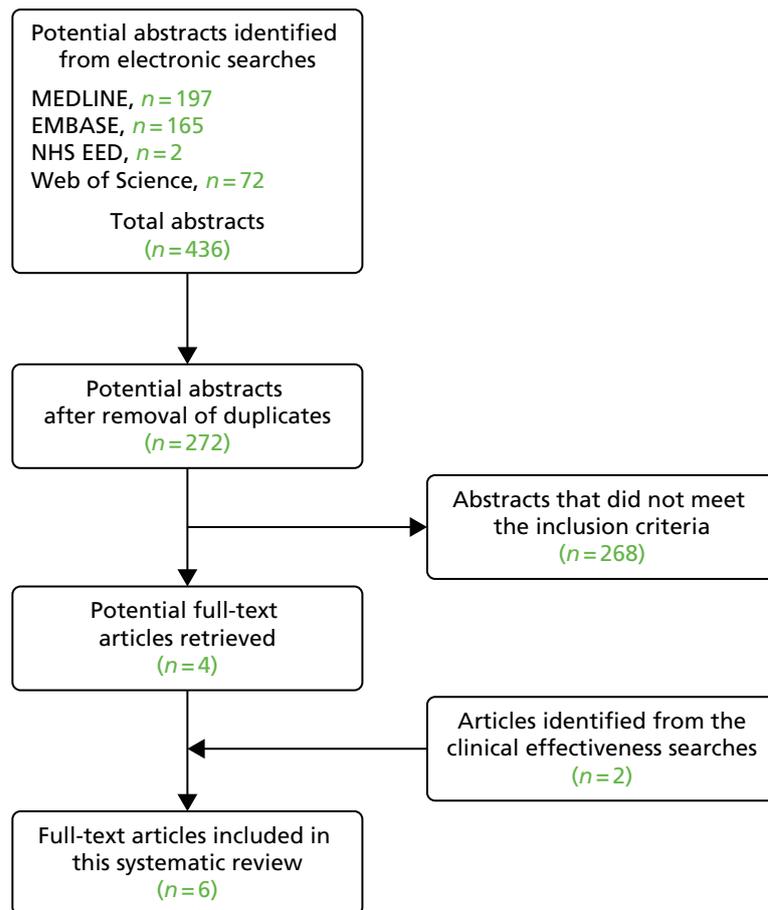


FIGURE 15 A PRISMA flow diagram for cost-effectiveness studies.

The HTA review by Clar *et al.*³ compared ACI with MF and mosaicplasty. and the authors attempted to calculate reliable costs per QALY. However, they felt that this was not possible because of the absence of data which were required. For example, QoL data were limited to around 2 years and no long-term studies (i.e. 20–30 years) were available on the incidence of OA and the need for TKR. The short-term modelling (QoL improvements at 2 years) found that the gain from ACI vs. MF would have to be between 70% and 100% greater over 2 years for the cost per QALY for ACI compared with MF to be within the £20–30,000 threshold. For the medium-term modelling (using 10-year success rates), the authors found that if the QoL gains were to be maintained for the next 10 years then for ACI relative to MF the QoL gain would have to be between only 10% and 20% greater to justify the additional cost of the intervention and be cost-effective within the £20–30,000 threshold. For the longer-term modelling there may be a need to offer some (or all) patients TKR, so a 50-year time horizon was considered appropriate. The authors found that for this scenario mosaicplasty was dominated, and moving from MF to ACI was associated with an incremental cost-effectiveness ratio (ICER) of between £3500 and £5500 (cells were assumed to cost only £3200). Overall, the authors concluded that there was insufficient evidence at the moment to say that ACI was cost-effective compared with MF or mosaicplasty.

The MSAC report³² compared the costs of MACI/ACI with mosaicplasty and MF in patients aged between 15 and 55 years suffering from a focal defect in an otherwise normal knee. In the absence of conclusive effectiveness data, MSAC assumed that the clinical effectiveness for all the different interventions was identical, and a cost-minimisation analysis was conducted. Resource use was determined by an Advisory Panel and the costs of the different procedures were obtained from various sources; for example, the cost of autologous chondrocyte transplantation was obtained from the prosthesis price list. The authors assumed that assessment costs and rehabilitation costs were identical, so they were not included in the

comparison. The price year (and time horizon) was not explicitly stated for the different resource use items, except for the prostheses (August 2010). The cost analysis found that the total costs of MACI/ACI (biopsy and grafting) procedure were significantly higher per knee than either mosaicplasty and MF (US\$14,083 vs. US\$2639 and US\$1405, respectively). The main cost difference between the procedures was that MACI/ACI required the cost of the chondrocyte cell culture and sealant (US\$11,780). MSAC felt that the conclusions that can be drawn from this review were limited by the quantity and the quality of evidence.

The updated MSAC report published in 2012¹⁴⁴ concluded that MACI was superior to MF (and mosaicplasty) with respect to less need for subsequent surgery and also in terms of clinical outcomes; therefore, a costing analysis was no longer sufficient and a cost-effectiveness/cost-utility analysis was required. A proposed model structure for the economic evaluation was presented using a decision tree with a Markov process, along with information on resource use and costs. They stated utility values would be obtained from the literature. However, no results of the cost-effectiveness analysis were presented.

Study design, intervention and patient populations

The remaining four peer-reviewed journal articles are summarised in *Table 12*. One study was a cross-sectional retrospective study (Derrett 2005¹⁴⁰) and the other three studies were decision-analytical modelling studies by Gerlier *et al.*,¹⁴¹ Samuelson *et al.*¹⁴² and Koerber *et al.*¹⁴³ The Derrett study¹⁴⁰ was conducted in the UK, and the other three studies^{141–143} were based on literature and some trial data from Belgium, Germany and the USA. Three studies assessed the cost-effectiveness of ACI compared with other interventions: mosaicplasty (Derrett 2005¹⁴⁰); MF (Gerlier 2010¹⁴¹); mosaicplasty and MF with different versions of ACI (ACI-C, ACI-P and MACI) (Koerber 2013¹⁴³). Samuelson *et al.*¹⁴² compared ACI-C with ACI-P to see whether or not it was more cost-effective.

The patient populations varied. The retrospective study by Derrett *et al.*¹⁴⁰ was based on 95 patients, of whom 53 patients received ACI, 20 patients received mosaicplasty and 22 patients were on the waiting list for ACI. The patients who received ACI were slightly younger than those who had received mosaicplasty (31.9 years vs. 34.9 years; $p = 0.17$) and more men received ACI (53% men vs. 47% women) compared with mosaicplasty (45% men vs. 55% women). The three economic models were based on clinical data and data from the literature.

Gerlier *et al.*¹⁴¹ compared adult patients who were less than 50 years of age (a mean age of 35 years at model entry), with symptomatic cartilage lesions of the femoral condyles, who had not yet developed OA, and the key efficacy data came from the TIG/ACT trial.⁴³

Samuelson *et al.*¹⁴² compared adult patients with a mean age of 30 years with a focal chondral injury which satisfied the conditions for an ACI repair.

The model by Koerber *et al.*¹⁴³ was said to be based on the model by Gerlier *et al.*¹⁴¹ In the Koerber supplementary file it was stated that the study population was patients aged 32 years with symptomatic, isolated cartilage defects and no contraindication. None of the economic models specified the number of hypothetical patients used for the modelling.

Time horizon and length of follow-up

The time horizon for any study should be long enough to capture all of the benefits that would accrue from the different interventions. The follow-up length in the studies varied. The Derrett study¹⁴⁰ was based on follow-up data for 2 years. The economic model by Gerlier *et al.*¹⁴¹ used two time horizons: a short-term time horizon of 5 years to take into account knee pain and mobility after the initial intervention (this information was obtained from a 5-year RCT that compared ACI with ChondroCelect and MF) and a long-term time horizon of 40 years to take into account the development of OA after 15 years and the need for a TKR after 20 years. Samuelson *et al.*¹⁴² based their model on a 10-year time horizon that corresponded with the longest-term evidence that was available in the literature. Koerber *et al.*¹⁴³ stated that on the basis of the

TABLE 12 Characteristics of previous economic studies

| Author, publication year, country | Aims, study design and patient group | Economic evaluation type, model, perspective, and currency and price year | Costs and outcomes | Results |
|---|---|---|--|--|
| Derrett 2005 ¹⁴⁰ Country: UK | <p><i>Aim:</i> To assess costs and health status outcomes after ACI and mosaicplasty</p> <p><i>Study design:</i> Cross-sectional retrospective study</p> <p><i>Patient group and numbers:</i></p> <ul style="list-style-type: none"> • 53 ACI recipients • 20 mosaicplasty recipients • 22 ACI WL recipients <p><i>Mean age (% male):</i></p> <ul style="list-style-type: none"> • ACI: 31.9 (53%) • Mosaicplasty: 34.9 (45%) • ACI WL: n/a (59%) | <p><i>Type:</i> Cost-utility analysis</p> <p><i>Model:</i> None</p> <p><i>Perspective:</i> Not stated</p> <p><i>Currency and price year:</i> UK £, 2003–4 prices</p> <p><i>Time horizon:</i> 2 years</p> <p><i>Discounting:</i> None</p> | <p><i>Resource use and costs:</i></p> <p>Operations/treatments, arthroscopies, inpatient stay, day case and outpatient visits, MRI scans, histology and X-rays</p> <p><i>Outcomes:</i></p> <ul style="list-style-type: none"> • Modified Cincinnati knee rating system • Pain Disability Index • EQ-5D-3L used to calculate QALYs <p><i>Sensitivity analyses:</i> One-way</p> | <p><i>Outcomes – EQ-5D means:</i></p> <ul style="list-style-type: none"> • ACI = 0.64 • Mosaicplasty = 0.47 <p><i>Costs:</i></p> <ul style="list-style-type: none"> • ACI = £10,600 • Mosaicplasty = £7948 <p><i>ICER:</i></p> <ul style="list-style-type: none"> • £16,349 cost per QALY |
| Gerlier 2010 ¹⁴¹ Country: Belgium | <p><i>Aim:</i> To assess the cost-effectiveness of ACI with CC compared with MF</p> <p><i>Study design:</i> Decision tree model</p> <p><i>Patient group:</i> Adult patients < 50 years of age with symptomatic cartilage lesions of the femoral condyles who had not developed OA</p> | <p><i>Type:</i> Cost-utility analysis</p> <p><i>Model:</i> Decision tree</p> <p><i>Perspective:</i> Global health-care payer (public payer reimbursement plus possible patient co-payment)</p> <p><i>Currency and price year:</i> Euros (€), 2008 prices</p> <p><i>Time horizon:</i> 5 and 40 years</p> <p><i>Discounting:</i></p> <ul style="list-style-type: none"> • Costs: 3% • Effects: 1.5% | <p><i>Resource use and costs:</i> Reimbursed drugs, medical procedures including ACI with CC and MF, consultations, hospitalisations and FU</p> <p><i>Outcomes:</i></p> <ul style="list-style-type: none"> • KOOS • SF-36 collected from a RCT used to calculate QALYs <p><i>Sensitivity analyses:</i> One-way, two-way and probabilistic</p> | <p><i>Outcomes – QALY means:</i></p> <ul style="list-style-type: none"> • CC = 21.08 • MF = 19.79 <p><i>Costs:</i></p> <ul style="list-style-type: none"> • CC = €29,808 • MF = €9006 <p><i>ICER:</i></p> <ul style="list-style-type: none"> • €16,229 cost per QALY |

TABLE 12 Characteristics of previous economic studies (continued)

| Author, publication year, country | Aims, study design and patient group | Economic evaluation type, model, perspective, and currency and price year | Costs and outcomes | Results |
|---|--|--|--|---|
| Samuelson 2012 ¹⁴² Country: USA | <i>Aim:</i> To assess the cost-effectiveness of ACI-C vs. ACI-P <i>Study design:</i> Decision tree model <i>Patient group:</i> Adult patients (30 years of age) with a focal chondral injury that satisfies the conditions for ACI repair | <i>Type:</i> Cost–utility analysis <i>Model:</i> Decision tree <i>Perspective:</i> Not stated <i>Currency and price year:</i> US\$, price year not stated <i>Time horizon:</i> 10 years <i>Discounting:</i> <ul style="list-style-type: none"> Costs: 3% Effects: 3% | <i>Resource use and costs:</i> Initial consultation, FU visits, surgical costs, ACI, physiotherapy, medical equipment <i>Outcomes:</i> <ul style="list-style-type: none"> Lysholm knee score Utility values from literature used to calculate QALYs <i>Sensitivity analyses:</i> Threshold | <i>Outcomes:</i> <ul style="list-style-type: none"> ACI-C = not stated ACI-P = not stated <i>Costs (total):</i> <ul style="list-style-type: none"> ACI-C = \$66,940 ACI-P = \$66,752 <i>ICER:</i> <ul style="list-style-type: none"> Not calculated |
| Koerber 2013 ¹⁴³ Country: Germany | <i>Aim:</i> To assess cost-effectiveness of mosaicplasty, ACI-P, ACI-C, MACI compared with MF <i>Study design:</i> Decision tree model <i>Patient group:</i> Patients aged 32 years with symptomatic, isolated cartilage defects and no contraindication | <i>Type:</i> Cost–utility analysis <i>Model:</i> Decision tree <i>Perspective:</i> German statutory health insurance <i>Currency and price year:</i> Euros (€), price year not stated <i>Time horizon:</i> 47 years <i>Discounting:</i> <ul style="list-style-type: none"> Costs: 3% Effects: 3% | <i>Resource use and costs:</i> Surgical treatments, inpatient stays, outpatient visits, arthroscopy, revisions, GP visits, imaging, physiotherapy and medications <i>Outcomes:</i> <ul style="list-style-type: none"> Utility values from literature used to calculate QALYs <i>Sensitivity analyses:</i> Probabilistic | <i>Outcomes – QALY means:</i> <ul style="list-style-type: none"> MF = 19.66 Mosaicplasty = 19.47 ACI-P = 19.76 ACI-C = 19.79 MACI = 19.80 <i>Costs:</i> <ul style="list-style-type: none"> MF = €13,445 Mosaicplasty = €17,774 ACI-P = €19,082 ACI-C = €18,713 MACI = €21,204 <i>ICER:</i> Cost per QALY gained in relation to MF <ul style="list-style-type: none"> Mosaicplasty is dominated by MF ACI-P = €56,370 per QALY gained ACI-C = €40,523 per QALY gained MACI = €55,421 per QALY gained |

CC, ChondroCelect; FU, follow-up; GP, general practitioner; n/a, not available; WL, waiting list.

German life expectancy of the patients in the model the time frame was set to 47 years. Although the authors did not explicitly state the cycle length – from the information provided this can be deduced as 1 year. Neither Gerlier *et al.*¹⁴¹ nor Samuelson *et al.*¹⁴² reported the cycle length that was used in the model, and none of the three studies applied a half-cycle correction to the economic models.

Study perspective and outcomes

Study perspective is crucial to the economic evaluation, as it will determine whether the appropriate resource use and costs have been collected, calculated and reported. Only two studies explicitly stated the viewpoint for the economic analysis: Gerlier *et al.*¹⁴¹ conducted the study from the perspective of the global health-care payer, whereas Koerber *et al.*¹⁴³ conducted their study from the viewpoint of the German statutory health insurance. All four studies conducted a cost–utility analysis for which the final outcomes were reported as QALYs. In addition, the Derrett study¹⁴⁰ used a range of outcome measures to compare the groups after surgery. The postoperative group consisted of patients who received either ACI or mosaicplasty who were compared with the ACI waiting list group. Outcome measures included:

- The Cincinnati knee rating scale, which assesses 11 components, including subjective symptoms such as pain and swelling, and functional activity level such as walking and climbing stairs scores – these scores were higher in the combined surgery group than the waiting list group.
- The Pain Disability Index, which helps patients measure the degree to which their daily lives are disrupted by pain – the authors found that patients in the combined surgery group had less pain than the waiting list group ($p = 0.09$).
- The generic health-related QoL – EuroQoL-5 Dimensions, three-level version (EQ-5D-3L) – measure. Patients in the combined group had statistically significantly higher EQ-5D scores than the waiting list group (0.61 vs. 0.41; $p = 0.03$). The EuroQoL-5 Dimensions (EQ-5D) measure was used to calculate the QALYs.

The Gerlier study¹⁴¹ used data from the SF-36 measure to calculate QALYs (this information was collected over a period of 60 months after randomisation from a RCT); in addition, they also used the KOOS, which evaluates five key dimensions: pain, symptoms, ADL, sport and recreation function, and knee-related QoL. Samuelson obtained utility values from the literature to calculate QALYs, although they did not specifically state which instrument or what method was used to estimate these utility values that were used in the model. In addition, some studies used in the model had used the Lysholm knee score (this measure contains eight domains, with a higher score indicating a better outcome) to estimate the utility values. Koerber *et al.*¹⁴³ obtained from the literature (no information sources were provided) and were based on the following: utility after treatment pain free (high functionality), utility with low functionality of the knee (medium functionality) and utility before knee prosthesis with strong pain (low functionality).

Resource use, costs and discounting

Derrett *et al.*¹⁴⁰ provided a comprehensive breakdown of resource use and costs, which were collected for the economic evaluation. These included secondary-care resource use related to each procedure, which was collected from patients' electronic and medical records from the time point of the first preoperative outpatient appointment to 2 years postoperatively. In addition, they also stated price year for which the costing was undertaken (year 2003–4). The resource use and costs of the surgical procedures, and the follow-up costs after initial interventions that were used in the model, have been comprehensively listed by Gerlier *et al.*¹⁴¹ This included information detailing the length of stay for each procedure and follow-up stage and also stated the price year for the economic analysis (year 2008). Both the Samuelson¹⁴² and Koerber¹⁴³ studies provided resource use and cost information, but not as detailed as the Derrett¹⁴⁰ and Gerlier¹⁴¹ studies. For example, for the different procedures the components were not individually listed and the price years for which the economic analysis was not explicitly stated – therefore researchers cannot use these unit costs for their own studies or to conduct a cost comparison with their own or with other studies.

All three of the economic models performed discounting using both 3% for costs and outcomes, except for Gerlier *et al.*,¹⁴¹ who used 1.5% for outcomes. Derrett *et al.*¹⁴⁰ did not conduct discounting, stating 'that costs tended to occur in the first year, making discounting unnecessary . . . the exact timing of postoperative benefit accrual was unknown'. Discounting is important in cost-effectiveness analyses, as it converts future costs into present values, thereby allowing comparisons between costs and benefits that occur at different times. This is especially important for different interventions for which costs usually occur in the current time period, whereas benefits are usually not evident until some point in the future; hence, discounting should have been undertaken by Derrett *et al.*¹⁴⁰ because the study length was greater than 1 year.

Base-case results and sensitivity analyses

The results and the conclusions offered by each study differed: Derrett *et al.*¹⁴⁰ found that the average cost was higher for ACI than mosaicplasty (£10,600 vs. £7948 in 2003–4 prices). Outcomes in terms of EQ-5D were better for the ACI group than with mosaicplasty (0.64 vs. 0.47); this difference was not statistically significant. Overall, the ICER for providing ACI relative to mosaicplasty was £16,349.

Gerlier *et al.*¹⁴¹ found that the mean costs of ChondroSelect ACI were higher than MF (€29,808 vs. €9006 in 2008 prices), but the overall mean QALYs were also higher for the ACI group (21.08 vs. 19.79). The authors found that the probability of ACI being cost-effective was approximately 80% if the payer has a willingness to pay of €22,000 per QALY. The cost per QALY gained for ACI over MF was €16,229.

Samuelson *et al.*¹⁴² found that the total costs of ACI-C were slightly higher than ACI-P – a difference of US\$188 (US\$66,940 vs. US\$66,752). However, there was some conflicting evidence when they later say that ACI-C was less expensive by US\$941. The earlier figure we presume relates to the initial cost difference and the latter figure must be after the model was run for 10 years; however, this was not explicitly stated. Also, no further information or breakdown was provided by the authors to show how these costs were obtained or calculated. Individual QALY means were not reported over the 10-year period, except the authors stated that ACI-C was more effective by 0.07 QALYs. The authors calculated a cost per QALY for each of the two different ACI interventions by dividing the cost of the intervention by the QALY to get a cost per QALY; however, this was not an incremental cost. Also, we could not work backwards to find out what these individual costs and QALYs were for each intervention. From the information gleaned from the paper, the ICER should have been reported as the cost per QALY gained of ACI-C relative to ACI-P – US\$13,443 (US\$941/0.07).

Koerber *et al.*¹⁴³ reported mean costs and QALYs for each intervention separately; the costs ranging from €13,445 (MF) to €21,204 (MACI), and QALYs ranging from 19.47 (mosaicplasty) to 19.80 (MACI). The cost per QALY gained was worked out for each intervention in relation to MF; the authors found that mosaicplasty was dominated by MF (MF was cheaper and more effective), whereas the cost per QALY gained ranged from €40,523 for ACI-C to €56,370 for ACI-P both in relation to MF.

Sensitivity analyses are important in economic analyses, as they deal with uncertainty around key parameters and assumptions made in the model, and help confirm the robustness of the results. All four studies conducted some sort of sensitivity analyses (SAs), ranging from the most simplistic one-way SAs (Derrett¹⁴⁰) to the more sophisticated probabilistic analyses (Gerlier 2010,¹⁴¹ Koerber 2013¹⁴³).

Limitations of the studies

All four articles had some methodological shortcomings. For example, in the Derrett study,¹⁴⁰ patients were not randomly assigned to treatment groups and follow-up was for only 2 years. The perspective of the economic analysis was not stated, and both costs and benefits were not discounted; only one SA was carried out, which looked at the lowering the costs of the ACI (where the ICER decreased slightly), and there were no preoperative utility scores for both groups (therefore utility values from a waiting list group

were used). Gerlier *et al.*¹⁴¹ felt that there were not enough data on the probability and time to occurrence for specific events such as TKR, which meant that a Markov model could not be developed. Another key limitation was the lack of long-term clinical follow-up data that could be used in the model, but one of the strengths of the study was the use of the data from the RCT to help populate the model. The limitations in the Samuelson study¹⁴² are most notably the inability to calculate the ICER (cost per QALY gained of ACI-C relative to ACI-P) accurately; short follow-up (10 years); perspective of the economic analysis was not stated; lack of trial data and the model relied heavily on assumptions and data from different studies in the literature; and lack of data on the QoL, that is the authors assumed utility values after both ACI-C and ACI-P were the same, as were the failure rates. Koerber *et al.*¹⁴³ did not explicitly evaluate ACI, but merely used ACI as an example to explain early evaluation and value-based pricing of regenerative medical technologies, although they did provide a supplementary file with some of the model inputs.

Quality assessment

The quality of the reporting of the economic analyses by the four articles was assessed using the 27-point CHEERS checklist.¹³⁸ Koerber *et al.*¹⁴³ did not identify the study as an economic evaluation in the title, nor did they provide a structured abstract. Only two studies reported the viewpoint of the economic analysis (Gerlier 2010,¹⁴¹ Koerber 2013¹⁴³). Samuelson *et al.*¹⁴² did not describe all of the comparators fully. The choice of health outcomes was well reported by all four studies; in terms of analytical methods and study parameters, these were best reported by Derrett *et al.*¹⁴⁰ and Gerlier *et al.*¹⁴¹ The article by Gerlier *et al.*¹⁴¹ was the most comprehensively completed in terms of economic analysis using the CHEERS checklist: 18 of the 27 statements (66.7%) were a 'yes', four statements (14.8%) were partially completed, two statements (7.4%) were not completed and three statements (11.1%) did not apply. The least comprehensive article in terms of the economic analysis was the article by Koerber *et al.*,¹⁴³ in which their study resulted in 'yes' to only 7 out of the 27 statements (25.9%); eight statements were partially completed (29.6%), five statements were not completed (18.5%) and three statements did not apply (11.1%).

Using the adapted Phillips *et al.*¹³⁹ 32-point checklist to critically appraise the economic models, overall the four articles adequately reported the objective of the model evaluation, the structure of the model, the type of model for the decision problem, the methods and assumptions to extrapolate short-term results into final outcomes, and the costs used in the model. The models did not provide clear justification if any feasible options were excluded. The cycle length was not explicitly stated in any of the studies, the choice of baseline data was not justified and none of the methods used expert opinion. None of the models applied a half-cycle correction, or justified that omission. Again, the Gerlier article¹⁴¹ was the most comprehensive analysis when using the Philips checklist¹³⁹ to critique the article: 21 out of the 32 statements (65.7%) were a 'yes', five statements (15.6%) were partially completed and six statements (18.8%) were not completed. The Samuelson article¹⁴² was not as comprehensively completed in terms of the economic model: only 8 of the 32 statements were a 'yes' (25.0%), 10 statements (31.3%) were only partially completed and nine statements were not completed (28.1%).

We also note an Austrian HTA report by Kunzl *et al.*¹⁴⁵ from the Ludwig Boltzmann Gesellschaft HTA unit, which commented that Austria was one of the few countries that funded ACI. However, the Ludwig Boltzmann Gesellschaft HTA report concluded that in 2009 there was a lack of evidence that ACI was more clinically effective than the other options. No cost-effectiveness analysis was performed.

Discussion

The cost-effectiveness search highlighted six studies that had been published since 2004; these studies were classed as full economic evaluations on the use of ACI, MF and mosaicplasty for repairing symptomatic articular cartilage defects of the knee. These studies included two technology assessment reports: one from the UK (Clar *et al.*³) and one from Australia (MSAC³²). In addition, there was one cross-sectional study from the UK and three economic modelling studies (one each from Belgium, Germany and the USA).

All the articles had shortcomings. The main limitations are summarised below:

- All models were decision models and no models were Markov-type models. A Markov model is more appropriate than a decision model because of the nature and progression of the disease and because articular cartilage defects can evolve over time.
- There was a lack of long-term clinical follow-up data and any studies with trial data were only for short periods (i.e. 2 years). The model would ideally need two time horizons: a short-term model (i.e. 3 years) to look at the short-term benefits of ACI and its comparators and a long-term model (i.e. 40 years) to look at the longer-term benefits of ACI and its comparators and the need for TKR.
- The models did not take into account all of the various health states that a patient with symptomatic articular cartilage defects of the knee can progress through over time.
- As all the economic models were decision models, transition probabilities were not reported. These probabilities are important for Markov models, as they show the direction and speed of transitions between the different health states.
- There was also a lack of good QoL data in each of the studies, and different instruments and methods that were used in estimating utilities/QALYs were not always reported. Good QoL data is important to show the benefits that evolve over time from ACI and its comparators.
- Finally, not all resource use, costs and price years were reported. Good resource use and cost data are important, as technologies are always evolving and accurate costings are needed to make comparisons with other treatments/interventions.

Chapter 5 Commentary on submissions by manufacturers and by the Oswestry group, including data from the ACTIVE trial

ChondroCelect

The submission on ChondroCelect was prepared by SoBi on behalf of TiGenix. ChondroCelect was developed by TiGenix, a cell therapy development company based in Belgium (www.tigenix.com). It was approved by the EMA in 2009, and the commercial launch in Europe was in 2010. The first country to approve reimbursement was Belgium in 2011, followed by The Netherlands in 2012. ChondroCelect was licensed to be marketed in Europe by SoBi (Stockholm) in 2014.

The submission starts with a concise and accurate account of cartilage structure and defects, and treatment options. It then goes on to present evidence of clinical effectiveness from four sources:

- The RCT TIG/ACT/01/2000.⁴³ (TIG is short for TiGenix.)
- A 'compassionate use' case series.
- A 'non-interventional' study – a registry-based cohort from routine care in Belgium and the Netherlands where ACI is funded, with 153 patients reaching 6 months or more of follow-up.
- The Belgian reimbursement scheme.

The submission notes the evolution of ACI over time. The TIG/ACT trial⁴³ used the Brittberg technique, using a periosteal flap (ACI-P). The compassionate case series used the same technique but with a collagen membrane (ACI-C). The manufacturer notes that current ACI mostly uses a cell-loading technique. The cells are loaded into the membrane by the surgeon.

As explained earlier, we regard ACI-P as now superseded because it requires more theatre time and has more subsequent costs (shaving of hypertrophy) but no clinical advantage.⁵⁷ However, we give details of the TIG/ACT trial⁴³ below. It was a good-quality trial but results may now be better, with ACI-C. We also give an account of the compassionate use case series and the other sources.

The product used in both trial and case series had 'characterised' chondrocytes.

Trial data: autologous chondrocyte implantation–periosteal flap versus microfracture: TIG/ACT/01/2000

This trial⁴³ compared ACI-P with CCI against MF in patients with symptomatic cartilage defects of the femoral condyles. The 5-year results are reported by Vanlauwe *et al.*⁴³ Other papers from this study include Saris 2008⁶⁹ and Saris 2009.⁷⁰ The former provides 12- and 18-month follow-up results and the latter has 36-month follow-up results.

Patient characteristics

Patients were aged between 18 and 50 years, with a single symptomatic cartilage lesion (ICRS grade 3 or 4) of size between 1 and 5 cm² in the femoral condyles of the knee, who agreed to follow a strict rehabilitation protocol.

A total of 118 patients were randomised: 57 to the ACI-P CCI group and 61 to the MF group. Six of the ACI patients were withdrawn because of failed chondrocyte expansion ($n = 1$) or negative ChondroCelect score ($n = 5$). ChondroCelect score helps predict whether or not the cells can grow into stable hyaline cartilage in

vivo. So only 51 patients were included in the ACI analysis. Details of baseline characteristics of these patients are from the Saris *et al.* papers.^{69,70} The mean ages of patients were similar in both groups (33.9 years). Most patients were male (61% in ACI and 67% in MF). Mean weights were similar (78.3 kg in ACI, 80.6 kg in MF, BMI not reported). Median durations of symptoms were similar (1.97 years in ACI, 1.57 in MF) and 37% in ACI and 21% in MF had had more than two previous knee procedures. In the ACI group, five had had previous MF, three had had subchondral drilling, and one had had abrasion arthroplasty. Only 12% of patients in ACI and 23% in MF groups had no history of previous knee surgery, including arthroscopy. At baseline arthroscopy, 98% of patients in ACI and 97% in MF had a single cartilage lesion, mostly grade 4 lesions. Mean sizes of defects after debridement were 2.6 cm² in ACI and 2.4 cm² in MF.

Patients in each group were categorised into re-intervention group (RIG) or no re-intervention group (NRIG) based on whether or not they underwent re-intervention on the index lesion during the study period. Seven patients in the ACI group and 10 patients in the MF group underwent re-intervention on their index lesion mainly because of recurring pain. In the ACI group, 5% patients in the NRIG group and none in the RIG group had BMI of > 30 kg/m².

Details of intervention and comparators

Details of intervention and comparators were given in the Saris *et al.* papers.^{69,70} All patients underwent baseline arthroscopy to assess eligibility to participate in the study. Patients in the MF group were treated following a technique recommended by Steadman *et al.*¹⁰¹ and those allocated to the ACI group were treated following the method recommended by Brittberg *et al.*²¹ Patients who were allocated to ACI group had cells implanted about 27 days after initial arthroscopy, secured beneath a periosteal flap.

Patients from both groups underwent an identical rehabilitation programme. In the first 2 weeks after surgery, patients were not allowed to bear any weight on their operated knee. After this, they were allowed to bear weight of up to 10–15 kg in the third week, and in the fourth to sixth weeks the weight was increased up to 15–30 kg. Then, the weight was increased progressively, as long as patients could tolerate it. For the first 8 weeks, all patients wore an unloader brace.

Duration of follow-up

Patients were followed up for 60 months.

Outcomes

At 12 months, cartilage biopsies were collected via arthroscopy from the middle of the repaired tissue for histopathological analysis. The primary outcome measure was change in overall KOOS from baseline at 36 months and 60 months. Other outcomes included AEs, changes from baseline in different KOOS domains, and analysis of overall KOOS after adjusting for the baseline covariates age, associated lesions and lesion location. Exploratory analysis was undertaken according to the time since onset of symptoms (< 3 years or ≥ 3 years) and age (< 35 years vs. ≥ 35 years).

Treatment failure was defined as 'a reintervention affecting more than 20% of the index lesion'. Time to treatment failure (TTF) was defined as 'the time between the end of the surgical procedure and the date of failure or reintervention'. All treated patients were included in the efficacy and safety population.

Results

Knee injury and OA outcome results were available from 43 patients in the ACI group and 45 patients in the MF group at both 36 and 60 months (*Table 13*). (To recap, an increase in KOOS indicates improvement. A score of 100 indicates no symptoms, a score of 0 is worst possible.)

At 60 months' follow-up, the overall KOOS and its subdomains improved in both treatment groups (see *Table 13*). The difference between the two groups was not statistically significant (7.10 95% CI -0.52 to 14.73; $p = 0.068$). In both treatment groups, the improvement in mean KOOS started as early as 6 months and was maintained up to 60 months' follow-up (*Table 14*).

TABLE 13 Mean change in overall KOOS and subscales^a from baseline at 60 months

| Score | 60 months (total group) | | |
|--|-------------------------|--------------|--------------------------------------|
| | ACI (SE) | MF (SE) | Difference (95% CI; <i>p</i> -value) |
| Overall KOOS | 21.17 (2.88) | 14.07 (2.54) | 7.10 (–0.52 to 14.73; 0.068) |
| ADL | 16.42 (2.97) | 11.35 (2.62) | 5.07 (–2.79 to 12.94; 0.203) |
| Pain | 19.04 (3.17) | 13.27 (2.74) | 5.77 (–2.55 to 14.09; 0.172) |
| Symptoms/stiffness | 17.70 (2.82) | 10.90 (2.52) | 6.81 (–0.70 to 14.32; 0.075) |
| QoL | 32.12 (4.30) | 21.23 (3.87) | 10.89 (–0.59 to 22.38; 0.062) |
| Function, sports and recreational activities | 32.50 (5.88) | 22.98 (5.69) | 9.52 (–6.87 to 25.90; 0.250) |

SE, standard error.

^a All subscales range from 0 to 100.**TABLE 14** Mean change in overall KOOS from baseline

| Time point | ACI | MF |
|----------------------|-------|-------|
| Baseline | 56.30 | 59.53 |
| Change from baseline | | |
| 6 months | 14.27 | 13.18 |
| 12 months | 16.96 | 13.54 |
| 18 months | 18.45 | 15.5 |
| 24 months | 19.38 | 13.09 |
| 30 months | 20.71 | 15.16 |
| 36 months | 21.35 | 14.72 |
| 60 months | 21.17 | 14.07 |

In the subgroup analysis according to the duration of onset of symptoms, the mean improvement in KOOS was greater in the ACI group than MF in patients with onset of symptoms of < 3 years' duration [25.96 (standard error, SE 3.45) vs. 15.28 (SE 3.17); difference 10.69, 95% CI 1.30 to 20.07; *p* = 0.026]. There was no significant difference in the mean KOOS between the groups in patients with onset of symptoms of > 3 years' duration [ACI 13.09 (SE 4.78) vs. MF 17.02 (SE 4.50); *p* = 0.554].

However, the groups differed in more than duration. Several factors that tend to lead to poorer outcomes were more common in the > 3 years group, including a higher proportion of previous lesions, fewer with acute onsets, and abnormal fellow knees.

Subgroup analysis by age found no statistical difference between the treatment groups [younger age patients: < 35 years ACI 22.4 (SE 3.70) vs. MF 16.59 (SE 3.55); *p* = 0.262; patient aged 35 years and more: ACI 19.61 (SE 4.51) vs. MF 15.16 (SE 4.01); *p* = 0.465].

Seven patients (13.7%) in the ACI group and 10 patients (16.4%) in the MF group had to undergo revision surgery on the index lesion. Most of the failures in the MF group occurred in the first 3 years, whereas those in the ACI group occurred in the fourth year or later.

The number of failures was lower in male patients than in female patients (ACI 6/19 females vs. 1/32 males, RR 4.21, 95% CI 1.03 to 17.57; MF 7/20 females vs. 3/41 males, RR 4.78 95% CI 1.49 to 15.62).

Radiographic results of 49 patients taken at baseline and at 60 months were available. The Kellgren–Lawrence grading of severity of knee OA showed no difference between the treatment groups at 60 months.

More patients in the ACI group experienced at least one related AE than those in the MF group (82% vs. 62%) during the 5 years. The AEs were mild to moderate in intensity. The most common AE reported was arthralgia (75% ACI vs. 62% MF in first 3 years; 36–60 months, ACI 14% vs. 4% MF). Other AEs included joint swelling (22% in ACI and 7% in the MF group in first 3 years; from 36–60 months, 0% in CCI and 2% in MF group), joint effusion (12% in ACI vs. 2% in MF between 36 and 60 months). None of the effusions was categorised as severe.

There were three AEs classed as serious in the ACI group and considered related to treatment: one DVT, one arthralgia and one tendinitis.

At the end of the follow-up, most of the AEs had disappeared, but there were 3 out of 37 cases and 1 out of 40 cases of effusion in the ACI and MF groups, respectively.

Commentary

Better results were seen with ACI in patients with shorter duration (< 3 years) of chondral defects.

Case series

The baseline characteristics of patients in the case series were more varied in some ways than in the RCT, as shown in *Table 15*.

TABLE 15 Comparison of baseline characteristics trial and case series patients

| Characteristic | RCT | Case series |
|-----------------------------------|--|---|
| Age (years), mean (range) | 34 (18–50) | 34 (range not given) |
| Male (%) | 64 | 57 |
| Duration of injury, years | Median 1.57, range 0–18 | |
| Site | Femoral condyles | Medial condyle 43% Patella 19% Lateral condyle 15% Trochlea 9% Condyle unspecified 7% |
| Previous procedures | 88% in ACI group, with 37% having had 2 or more, 'in particular marrow stimulation' | NR |
| BMI of > 30 kg/m ² | 10% | NR |
| Mean BMI (kg/m ²) | NR | 25 |
| Mean weight (kg) | 81 | NR |
| Inclusions | Symptomatic single lesion of femoral condyles, between 1 and 5 cm ² in size | No predefined entry criteria. |
| Exclusions | Significant knee abnormalities, patellar lesions, OA, previous mosaicplasty, MF in previous year | Active infection at biopsy site, significant OA, drug allergies |
| Size of lesion (cm ²) | 1–5 | 3.5 (0.2–20) |
| NR, not reported. | | |

The outcomes in the compassionate case series were the Clinical Global Impression measures of improvement (CGI-I) and efficacy (CGI-E). CGI-I measures change from baseline (no change, improvement, worsening). CGI-E has four points: very good, moderate, slight and no change or worse. Results were divided into short-term follow-up (under 18 months, mean 9 months, which is too short for best outcome) and longer term (> 18 months, mean 27 months) but figures in each group are not given.

Note that these scales are reported by the surgeon, not the patient. The CGI-I results were reported as showing good outcomes (much improved or very much improved) in 68%, with serious worsening in only 2%. The CGI-E results showed 38% with very good results, 36% with moderate improvement, 12% with slight improvement and 11% unchanged or worse. (From submission table 10, p. 30 – results total 97% not 100%.)

The submission reports that no differences were seen by duration of follow-up (< 18 months vs. > 18 months), site of lesion (patella vs. condyles) or size of lesion (small vs. large, not defined). Patients with single lesions did better than those with multiple ones, but only significantly so in CGI-I results (improved 86% vs. 77%). Results in multiple lesions were good.

The most common AE was knee pain (24%) and 54% had no AEs. As expected with the ACI-C method, few patients (2%) developed cartilage hypertrophy.

Registry cohort

Details from this cohort are sparse and only about half of the cohort (153 out of 308) had 6 months or more of follow-up. The mean age of 32 years (range 15–50 years) is similar to the RCT and case series. The only benefit reported is an increase in KOOS, at up to 36 months, but numbers at each follow-up period are not given. There were six treatment failures, and two DVTs among a total of 17 SAEs (but no denominator given). Treatment failure was defined as the need for a re-intervention for more than 20% of the treated area, associated with symptoms. The summary states that no new AEs were reported in the registry cohort.

Belgian reimbursement scheme

Little information is reported from this source. Two procedures failed within 12 months and another two procedures failed between 12 and 24 months, in 254 patients. Only 51 patients had reached 3 years of follow-up.

The data show an increase in numbers treated, from 51 in the first year (May 2011 to April 2012), 93 in the second and 110 in the third, possibly suggesting levelling off in numbers. The population of Belgium is 11.2 million, so the third year rate is about 10 per million per year. The equivalent numbers per year in England would be 540, and in Wales 30.

The ChondroCelect submission argues, with some justification (see *Chapter 2* of this report), that ACI is more successful as a primary procedure in patients who have had previous MF. The Minas *et al.* study⁷⁵ is cited in support of the assertion.

Cost-effectiveness

The health-related QoL was measured in the TIG/ACT trial⁴³ using the SF-36 administered at 18, 24, 30, 36, 48 and 60 months post procedure. At 36 months, SF-36 scores were slightly better for ACI.

Introduction and model structure

The economic analysis by SoBi used a de novo Markov model to assess the cost-effectiveness of ACI in relation to MF from a NHS and Personal Social Services (PSS) perspective. Both costs and outcomes were discounted at a rate of 3.5% per annum in line with NICE guidelines. Only the written assessment was provided to the Assessment Group. The model used is simpler than the Warwick model, but is regarded as fit for purpose.

Microfracture was considered to be the only relevant comparator for ACI, and other comparators such as mosaicplasty were not considered for this analysis – this is a reasonable assumption. The SoBi submission states that mosaicplasty is little used and ‘not recommended by NICE’. The last assertion is not quite correct. The NICE Interventional Procedures Guidance (2006),³⁴ which is concerned only with safety and efficacy (not cost-effectiveness), states that there were no major safety concerns and mentions ‘some evidence of short-term efficacy but data on long-term efficacy inadequate’.

Evidence of benefit came from a RCT with only 1 year of follow-up, in which ACI was better, and from case series with 2 or 3 years’ follow-up. NICE recommended that mosaicplasty should be used only with ‘special arrangements for consent and audit or research’. So it is correct to say that NICE has not recommended mosaicplasty in routine care.

The SoBi model is similar to the Warwick Assessment Group model, whereby patients enter the model at the time that they receive the procedure (ACI or MF). However, there are differences between the Warwick model and the SoBi submission: the cycle length used in the submission model is 1 month, whereas the Warwick model used a cycle length of 1 year. The average age of patients receiving a procedure in the SoBi model is 33 years and the model has time horizon of 75 years: on this basis the model assumes that patients can live up to an average age of 108 years (however, they did state that by this point > 99.9% of patients will have died). The model is separated by gender, but we know that there is no difference in the success or failure of the two different procedures if lesions are comparable.¹⁴⁶

The model structure is logical and similar to the Warwick model, as it allows both temporary and permanent successes. If either MF or ACI fail, the patient has debridement to remove the damaged tissue and can go on to receive another repair, but this second repair is MF only. Otherwise the patient may choose not to have a repair and is offered conservative pain relief treatment only. If this second repair (MF) fails, the patient will receive debridement and pain relief only.

Patients who receive best supportive care (BSC) may deteriorate and are assessed for a TKR. The SoBi model assumes that a patient can receive up to a maximum of only three TKRs. The modelling uses TTF as the outcome that drives the ICERs, using 5-year data from the TIG/ACI RCT and the case series. Delaying treatment failure leads to postponement of TKR costs. If the second TKR fails then the patient receives just analgesics. The following is unclear from the SoBi model:

- The average age that a patient will require a TKR.
- As evidence has shown, some patients may receive more than two TKRs.
- Also, the first knee replacement can either be partial or a TKR. As described later, this affects the costs of the second replacement.

Finally, the model assumes that patients can die at any stage from all-cause mortality, and there is a low risk of mortality from undergoing a TKR or a TKR revision.

Model inputs

Efficacy of first treatment

The SoBi model uses TTF as a proxy measure of treatment efficacy (i.e. when a new procedure for the same defect was required). This information on TTF (i.e. transition probability for moving from primary treatment success to treatment failure) was obtained from KM plots as reported in the Vanlauwe *et al.*⁴³ article. This article⁴³ reported that ACI was better than MF, and that patients in the ACI group waited longer before needing a further procedure because of the longer benefits. This is a reasonable assumption for the model.

Four different scenarios were used for the TTF after the observed data: scenarios 1–3 assumed no ACI benefit after the observed data, or after 10 or 20 years, at which point then the benefit of MF is applied to

the patient cohort; scenario 4 used the line of best fit for the entire model duration. For all scenarios, ACI was better than MF. These scenarios seem plausible.

Another four different scenarios were also used for treatment failure using observational ACI data (to take into account a normal clinical setting rather than a trial setting). The observed failure rates for ACI were 0.79%, 1.39% and 0.00% in years 1, 2 and 3, respectively. A weighted average failure was calculated as 0.89% and this was applied. Scenarios 1–3 assumed no ACI benefit after the observed data, or after 10 or 20 years, and in scenario 4 it was assumed that the average ACI benefit was maintained.

Subsequent treatment

The SoBi model in the base-case analysis assumed, based on clinical advice, that, when ACI fails, 90% of the patients will receive MF, and, that, when MF fails, only 5% of patients receive another MF. As the manufacturers said, this latter value is too low (these values are set to 50% in the SA). The submission did not say why patients who receive a first MF are less likely to receive second MF than patients who receive an ACI first.

Two papers from the TIG/ACT trial⁴³ reported failure rates for subsequent MF: Vanlauwe *et al.*⁴³ reported MF failure rate of 16.4% at 5 years (converted monthly rate 0.30%) and Saris *et al.*⁷⁰ reported MF failure rate as 11.5% at 3 years (converted monthly rate 0.34%). The latter value was used in the SA. The submission assumed that, based on clinical advice, a second MF following a first MF would be half as effective, that is twice the failure rate.

Two studies that reported failure rates for debridement were used for BSC following initial and subsequent treatment failure in the analysis: Forster *et al.*¹⁴⁷ reported a failure rate of 20.0% at 1 year (converted monthly rate was 1.84%) and Bernard *et al.*¹⁴⁸ reported a failure rate of 18.0% at 5 years (converted monthly rate was 0.33%). The latter value was used in the SA. Failure of BSC leads to knee replacement.

For TKR, based on expert clinical advice, the SoBi model assumed that 95% of the cohort would be suitable for a TKR and that a TKR is expected to last for 10–20 years (a midpoint of 15 years was used in the base-case model and was converted into a monthly transition probability). For those patients who need a TKR revision, the model assumed that there was a slightly higher failure rate than for the first TKR, and the first TKR will only last for 10 years – these are plausible assumptions for this patient group.

Mortality

Swedish Orphan Biovitrum AB used Office for National Statistics (ONS) data for all-cause mortality (split by age and gender) and, for the base-case, TKR mortality data was based on a figure reported on the NHS Choices website¹⁴⁹ (1.6%). A paper by Mahomed *et al.*¹⁵⁰ was used for TKR mortality (0.7% for initial TKR and 1.1% for a revision TKR) in a SA. The SoBi model assumed that the mortality rate for TKR revision would be 2.5% (i.e. based on Mahomed *et al.*,¹⁵⁰ a 57.1% increased risk). This is a reasonable assumption, as this is a longer operation, patients are older and rehabilitation might be slower.

Costs

The costs for the different procedures, rehabilitation, TKR, TKR revisions and pain relief were obtained from UK sources, literature and the HTA report by Clar *et al.*³ The cost of procedures included the costs of surgery, inpatient stays and physiotherapy follow-up. The submission stated that cost of TKR could not be identified from the NHS reference costs so they used information from the previous HTA report³ (whereas the Warwick model uses a NHS reference cost¹⁵¹ for TKR). The costs that have been inflated from the previous HTA report by Clar *et al.*³ are underestimated, as the wrong base-case year was used: the submission model should have used the year 2003–4 prices instead of 2005–6 prices. The inflation multiplier will have been 1.286 instead of 1.200. For example, the cost of MF as an inpatient procedure should be £3020 instead of £2818.

The submission reports that 'All costs are updated to 2014 using the latest Hospital & Community Health Services (HCHS) index' – when, in fact, the prices are uplifted to year 2012–13. We have not amended any

of the costs below, as this would mean that the total costs and ICER value would be different. However, we believe that the magnitude and direction of the costs differences will not change substantially.

The cost of ACI included the cost of the product including two-way courier and cell culture (£16,000) plus the cost of arthroscopy and cell harvest (procedure 1: £722.45) and arthrotomy conducted in an outpatient setting (procedure 2: £109.65). However, the cost for implantation of the cells is an underestimate, as the procedure would be done as a day case, not an outpatient visit. The total cost of ACI was £16,832.10. Adjustment of the cost of the second procedure gives a total cost of ACI of £16,832 + £722 = £17,554. The SoBi model also included the cost of a TKR assessment, which included a general practitioner (GP) assessment and cost of an outpatient appointment (£146.65) whether the patient went on to receive a TKR or not. The costs for TKR and TKR revision (£6500.85 and £12,093.24, respectively) look correct.

The SoBi model also included the cost of rehabilitation after ACI, MF and TKR in line with the Warwick model. However, the cost used by SoBi is lower than the cost used in the Warwick model (£42.47 vs. £256.00). In addition, the submission model also included the cost of pain relief medication – which consisted of paracetamol (this cost was not included, as the patients would have purchased this over the counter) and NSAIDs. SoBi estimated a weighted average cost for NSAID per month as £9.79. This cost is negligible and has not been included as a cost in the Warwick model.

The SoBi model also included a cost for patients who were classed as ‘unresolved patients’. This cost was estimated at £384.43 per year, which included the cost of GP visits, treatment visits, medications, outpatient visits, physiotherapist, prescribed aids (not specified but presumably walking aids), and complementary (not specified) and other therapies. This total cost was based on patients with lower limb OA, but for some patients this cost may be an overestimate, as some of these patients may just have pain relief medication and choose to put up with the pain.

The different cost values were varied in the SA.

Health-related quality of life

The SoBi submission states that there is lack of utility data in patients with a knee cartilage defect. Utility scores were based upon analysis of the SF-36 questionnaires, which were collected up to 60 months post surgery, as reported in Gerlier *et al.*¹⁴¹ in table II. These are plausible utility values. The model also accounted for the decreasing utility over time by using age-related UK population EQ-5D weights as reported by Kind *et al.*¹⁵² The model assumed that after successful ACI and MF, patients would have the same benefits, and the utility value used after surgery was 0.8170. The model does not specifically state how long this benefit lasts but we assume it is 5 years in line with the Gerlier *et al.*¹⁴¹ paper. This does not take into account that after MF the utility value will stay at this value for a few years but is likely to decline later, eventually to the presurgery value, as these patients are most likely to require another repair. Values were varied in the SA.

Adverse events

Adverse events were not included, as SoBi stated that there were no key differences between the two treatment arms.

Model results

The total cost of ACI was £22,586. The total cost of MF was £13,547. Total QALYs gained for ACI compared with MF were 1.29. The ICER for ACI compared with MF was £7077 per QALY. The main cost drivers were the cost of the cells and the fact that fewer people needed further repair or TKR with ACI than with MF. The model also assumed that further QALYs are gained by ACI patients when they received a subsequent MF (4.15 more QALYs when looking at QALY results disaggregated by health state) than MF patients when they received a subsequent MF (as these patients will fail more quickly).

The SoBi SAs found that the ICERs for the different efficacy scenarios and the subsequent treatment efficacy scenarios, as listed earlier, were consistent with the base-case analysis, that is, although ACI was

more expensive it was also more effective. For the subsequent treatment scenario in the base-case analysis, the use of subsequent MF after ACI was 90%, but only 5% had a second MF after the first MF (i.e. only a small proportion of patients would receive a second MF). In the SA this value was changed so that 50% would have a second MF after both ACI and MF. The resulting ICER was nearly £25,000. This is due to more people having MF and the QALY gain being lower (0.46 vs. 1.29).

The ICER was also sensitive to the model time horizon. For example, if a 5-year time horizon was used, the resulting ICER was approximately £290,000. This was due to the majority of costs of ACI being incurred in the first few years, and the benefits from ACI not being seen until later, in other words fewer people moving to an unresolved state and fewer people in need of a TKR. The model became cost-effective only when it was run for 20 years (ICER approximately £22,000). The ICER was robust to other scenarios that were tested, such as different utility values, TKR mortality and discounting. The probabilistic SA results were similar to the deterministic, with ACI probably the most cost-effective around the £6000–7000 range (i.e. a 98.8% chance of being cost-effective).

Overall, the model assumptions and results look plausible.

Aastrom Biosciences submission

Aastrom have now changed their name to Vericel Corporation.

The submission from Aastrom was based mainly on the SUMMIT trial,¹⁰⁰ including the extension study up to 3 years (it will, in time, produce 5-year data). The SUMMIT trial¹⁰⁰ was described in detail in *Chapter 2*.

Data from the Basad study⁸ were also presented.

The Aastrom submission states that an indirect comparison of MACI and ACI was performed, using MF as the common comparator, but this is illustrated by two separate forest plots: one showing the SUMMIT¹⁰⁰ results for MACI versus MF and the other showing the Saris *et al.* results⁷⁰ for ACI versus MF. Some RRs for SUMMIT¹⁰⁰ versus Saris *et al.*⁷⁰ are then presented, but the underlying methods and calculations are not provided. However, results were similar and CIs overlapped with 1. So no claim for clinical effectiveness superiority of MACI over ACI is made.

Data on ACI versus mosaicplasty are presented and used to argue, reasonably, that MACI is superior to mosaicplasty.

Aastrom argues that the main comparator is MF, particularly as the lesion sizes considered in the submission (3–20 cm²) are too large for mosaicplasty.

Cost-effectiveness

The submission by Aastrom Biosciences did not provide any cost-effectiveness analyses because of the recent purchase of the MACI product by Vericel from Sanofi. Cost-effectiveness evidence was presented in the MSAC submission³² and the manufacturers aimed to adapt this. However, this was not possible because of time constraints, so only a budget impact/costing forecast model was provided.

The budget impact model estimated by Aastrom indicated that 9549 patients in England and Wales were eligible for cartilage repair in 2013. Of these 9549 patients, as indicated by the NICE scope, 500 will be eligible for MACI or an ACI in year 5. The manufacturers assumed that there would be an equal split of the use of MACI and ACI. The rest of the eligible patients would receive MF, though the reasons for not offering ACI are not explained. Based on data from three studies (Minas 2009,⁷⁵ Nawaz 2014,⁸⁰ Vijayan 2014¹⁵³) the manufacturers reasonably assumed that reoperations after MF do not have the same success rate as primary MACIs or ACIs.

List prices were used for the costs for ACI (£18,300) and MACI (£16,226 excluding VAT). The cost of MF was £2464, which was obtained from the NHS reference costs.¹⁵¹ Cost of theatre, surgery for implantation of MACI/ACI was assumed to be the same as the cost of MF, though our clinical opinion is that MF usually requires an inpatient stay (because of pain), whereas ACI is usually a day case procedure. The Aastrom assumption may therefore slightly disadvantage MACI. The submission states that patients have one procedure. It is not clear whether this means that they would have only one MACI or it is an error by not accounting for both arthroscopy and harvesting, and later implantation. The manufacturers assumed that the cost of MACI/ACI reoperation would be £16,226. The cost of initial MF with MACI/ACI as second repair at an average cost of £17,623 also seems appropriate. This extra differential cost approximates to 3.5 extra rehabilitation visits. The cost of rehabilitation was £376, which was obtained from the Unit Costs of Health and Social Care (Curtis 2013¹⁵⁴). This cost was based on the use of a community-based therapist, with 6–8 rehabilitation sessions, each lasting 30 minutes. Alternative rehabilitation costs could have been obtained from the NHS reference costs.¹⁵¹ The budget impact model did not take into account any outpatient visits and any inpatient stays for MACI/ACI.

Three-year probabilities for MACI reoperation and MF reoperation were obtained from the SUMMIT trial data¹⁰⁰ and these were converted to annual probabilities: 0.005 and 0.014, respectively. The annual probability for MACI was also assumed for the ACI reoperation, which seems a reasonable assumption. The Saris *et al.* 2009⁷⁰ data provided alternative 3-year probabilities for reoperation – these were converted to annual probabilities: 0.010 for ACI/MACI reoperation and 0.040 for MF reoperation. The manufacturers assumed that if MACI/ACI failed then a reoperation would be either MACI/ACI; however, if MF failed then a reoperation would be MACI/ACI.

The budget impact model explored two scenarios: one scenario with MACI/ACI as first-line treatment and the other scenario without MACI/ACI (with MF only). Using failure rates based on the SUMMIT data¹⁰⁰ there were total cost savings from using MACI/ACI, ranging from approximately £5.9M in year 1 to £8.3M in year 5 – this was due to the lower reoperation rate and the expectation that 500 procedures (of the approximate 10,000 procedures) were either MACI/ACI. The Aastrom submission also included a budget impact model using the higher failure rates from Saris *et al.* 2009.⁷⁰ There were further total cost savings, although lower than the cost savings when the SUMMIT trial¹⁰⁰ failure rates were used – using MACI/ACI the cost savings ranged from approximately £5.8M in year 1 to £7.8M in year 5 – these lower cost savings were due to the need for more reoperations after MF. In conclusion, the cost calculations provided by Aastrom seem reasonable and plausible.

Submission by OsCell

The Oswestry submission included interim data from the ACTIVE trial,³⁵ which had about 5 years to run.

The ACTIVE trial

The ACTIVE trial³⁵ is a Medical Research Council-funded multicentre RCT of ACI against standard treatment, which could include debridement, abrasion, drilling, MF, mosaicplasty or bone graft (according to surgeon's discretion) in 390 patients (195 in each group) with symptomatic chondral defect(s) on the medial or lateral femoral condyle, trochlea or patella, who had failed previous treatment and who were also considered suitable for ACI/MACI.

The protocol is available online from the ACTIVE trial³⁵ website (www.active-trial.org.uk/).

Details of the results were made available to NICE for the ACI appraisal as academic in confidence. Only some of the results are reproduced here, with permission from the ACTIVE trial³⁵ group.

Patients were recruited from 29 centres. The RJAH in Oswestry recruited 87 patients (22%). Six centres recruited between 20 and 29 patients, six centres recruited between 10 and 19 patients, and 16 centres recruited fewer than 10 patients.

Quality assessment

Using the modified Coleman methodology score, the study scored a total of 73, suggesting that the quality is good.

The cells used came from two sources. In the Oswestry centre, the locally produced cells were used, but in all other centres commercially produced cells were used. So ACTIVE³⁵ is a trial of 'traditional ACI' only in the Oswestry centre.

The first primary outcome was to have been time to cessation of benefit, but this proved difficult to measure, and the second primary outcome, Lysholm assessor outcome score, was used. (The submission uses the phrase 'independently assessed'.) Other outcome measures included patient-assessed Lysholm score, Cincinnati knee score, IKDC score and EQ-5D.

As part of secondary outcomes, patients were asked to state their rating of operation at all follow-up points with responses ranging from extremely pleased to much worse (*Table 16*).

The Oswestry group report that ACI-C with Chondro-Gide™ (Geistlich, Wolhusen, Switzerland) is better than ACI-P, as the former leads to repair of the defect with better-quality tissue. The evidence for this comes from a study (McCarthy and Roberts¹⁵⁵) comparing the two in 88 Oswestry patients: 55 treated with ACI-P and 33 patients treated with ACI-C.

Cost-effectiveness

The cost-effectiveness analysis by Oswestry was based on the ACTIVE trial³⁵ data, but using cell costs only from Oswestry. The cells used in other ACTIVE³⁵ sites came mainly from Genzyme. The submission provided the costs of ACI and the different comparators. The benefits were in terms of QALYs, which were estimated from the EQ-5D-3L.

The total costs and incremental costs with and without the market forces factor (MFF) provided in the submission have been summarised in *Table 17* (MFF estimates the unavoidable cost differences of providing health care.) With Payment by Results (PbR), the MFF directly funds providers for the relative level of unavoidable costs they face. Each NHS Trust receives an individual MFF value used to establish the level of unavoidable costs they face relative to each other.

Accounting for unavoidable costs ensures a level basis across the country to provide equal amounts of health care per pound. So, in terms of PbR income, this would equal the activity multiplied by its tariff price, and this is then multiplied by the MFF value. All costs are in 2014–15 prices in UK pounds sterling (£).

TABLE 16 Patients' responses

| Time point (years) | Proportion (%) pleased or extremely pleased | |
|--------------------|---|----------|
| | ACI | Controls |
| 1 | 60 | 50 |
| 2 | 58 | 47 |
| 3 | 57 | 45 |
| 4 | 52 | 47 |
| 5 | 59 | 43 |

TABLE 17 Costs of ACI and its comparators by Oswestry

| Procedure | Costs (2014–15 prices, £) | Costs (£) including MFF | Incremental costs (£) of ACI over the comparator (including MFF) |
|---------------------------|---------------------------|-------------------------|--|
| Intervention – ACI | | | |
| First stage | 2398 | – | – |
| Second stage | 6876 | – | – |
| Total cost of ACI | 9274 | 9565 | – |
| Comparators | | | |
| TKR | 5642 | 5819 | 7094 |
| MF | 2396 | 2471 | 3746 |
| Osteotomy | 2396 | 2471 | 3746 |
| Mosaicplasty | 2396 | 2471 | 3746 |

The second stage for ACI includes the cost of the cells. Production of cells in Oswestry costs £4125 per patient, but this does not include overheads and set-up costs. The submission stated that the incremental cost of ACI over TKR was £3746 and the incremental cost of ACI over MF, osteotomy or mosaicplasty was £7094.

Autologous chondrocyte implantation costs included operations, hospital stays, the cells and any further implants. The TKR cost is in line with NHS reference costs (2012–13), which is £5676 (NHS reference costs 2012–13¹⁵¹). The costs included only the direct costs of the procedures. No information in the submission was provided on any further outpatient or rehabilitation visits. The submission also stated that further data have been collected using patient diaries on patient and societal costs, such as any out-of-pocket expenses and time off work, but were not available in time for this submission.

In the ACTIVE trial,³⁵ QALYs were estimated using the EQ-5D-3L version, weighted using the UK utility values. The control group consisted of patients who received MF, MF plus collagen membrane, and mosaicplasty. The ACI group had 192 observations at baseline, with a mean EQ-5D index score of 0.506, and by 8 years this score had increased to 0.676 ($n = 29$). The control group had 187 observations at baseline with a mean EQ-5D index score of 0.532 (slightly higher than the ACI group) and by 8 years this score was 0.647 ($n = 29$); this score was slightly lower than the corresponding score for the ACI group at 8 years. One year postoperatively, the EQ-5D index score had risen to 0.636 in the ACI group and 0.675 in the control group. The submission stated ‘these mean 1-year levels are largely maintained up to seven years post-surgery, after which the mean score for the control group seems to drop slightly’. Too few patients had reached the later years for a reliable analysis. Cumulative QALYs were also presented, which represented the total increase in health utility. These cumulative QALYs increased at a similar constant rate for both the ACI and control groups up to 6 years. The submission concluded that ‘at the 8-year time point, the ACI group has experienced – on average – an incremental gain of 1.19 QALYs more than the control group’ – this figure relates to cumulative QALYs and not mean QALYs.

Numbers of patients and EQ-5D for later years are reproduced in *Table 18*.

The submission concluded that the trial found that the ACI group gained more QALYs than the control group (incremental gain = 1.19 QALYs) at the 8-year time point, which resulted in a crude estimate of an incremental cost per QALY gain of £5961 for ACI over the control group.

In the base case, both groups were treated as homogeneous, but because of differences in the treatments for the control group and cell sources for the ACI group, further analyses tested for heterogeneity in

TABLE 18 The EQ-5D scores in the ACTIVE trial³⁵

| Time point, years | Standard arm | | ACI arm | |
|-------------------|-----------------|------------|-----------------|------------|
| | No. of patients | Mean EQ-5D | No. of patients | Mean EQ-5D |
| 5 | 87 | 0.618 | 102 | 0.677 |
| 6 | 75 | 0.675 | 80 | 0.707 |
| 7 | 53 | 0.705 | 59 | 0.735 |
| 8 | 27 | 0.647 | 29 | 0.676 |
| 9 | 5 | 0.486 | 2 | 0.761 |

each group. For the control groups (MF, MF plus collagen membrane and mosaicplasty), the data from the ACTIVE trial³⁵ suggested little difference in the EQ-5D scores. For the ACI group, as the cells came from different sources, the submission included a regression analysis to see whether or not cell origin might affect their benefit. The regression analysis provided a negative value from which the authors concluded that 'ACI patients treated outside Oswestry are unlikely to have more benefit from ACI'.

The absolute values for EQ-5D from the ACTIVE data³⁵ often show little difference, as shown in *Table 18*, but changes from baseline EQ-5D show a more consistent advantage for ACI, as shown in *Table 19*.

Discussion: clinical effectiveness

The four main trials^{8,35,43,100} described in this review all show some superiority of various forms of ACI over MF, but in different timescales and to different degrees. The Basad⁸ and SUMMIT¹⁰⁰ trials show clear differences in favour of MACI by 2 years in Lysholm and KOOS, respectively. The SUMMIT trial¹⁰⁰ found no significant difference in EQ-5D visual analogue scale (VAS) changes – both groups improved by 17 at 2 years. The TIG/ACT trial⁴³ shows superiority overall by 3 years. The ACTIVE trial³⁵ (in which ACI was a mixture of ACI-P, ACI-C and MACI) showed no benefit in most outcome measures at 5 years, but some separation in EQ-5D after that. With the exception of EQ-5D, results are available only to 5 years in an interim analysis provided for the NICE appraisal. The ACTIVE trial³⁵ will continue to 10-year follow-up.

Previous repairs

As reported earlier, in case series previous MF appears to reduce the success of ACI. The trials reviewed above do not contribute much evidence on this. Basad did not give details of previous surgery. In TIG/ACT⁴³ only a few (8/57) of the ACI group had had previous MF. In the SUMMIT trial,¹⁰⁰ 32% of the MACI

TABLE 19 Increases in EQ-5D from baseline

| Time point, years | Standard care | ACI |
|-------------------|---------------|-------|
| 1 | 0.143 | 0.130 |
| 2 | 0.135 | 0.163 |
| 3 | 0.141 | 0.168 |
| 4 | 0.142 | 0.141 |
| 5 | 0.086 | 0.171 |
| 6 | 0.143 | 0.175 |
| 7 | 0.173 | 0.229 |
| 8 | 0.115 | 0.170 |

group had had previous repair attempts with MF, but this appeared to have little effect on response rates (no prior repairs, 90% response; more than one, 84%). In the ACTIVE trial,³⁵ almost half had a previous repair procedure, but results are not given separately for them.

Several factors need to be considered in interpreting the evidence. First, we are somewhat reliant on subgroup analysis. Second, those who have had previous surgery may be older than those going straight to ACI, and chondrocyte viability declines with age. Third, some of the older trials had few patients who had not had prior surgery. Last, and most important, the evidence does not suggest that ACI is not worthwhile after prior MF, but only that it is not as successful. Hence there is no reason not to try ACI.

Duration of symptoms

In the SUMMIT trial,¹⁰⁰ response rates were similar at 2 years: 82% for those with symptoms for less than 3 years, 92% in those with longer durations. Basad did not report results by prior duration but his MACI patients had a mean duration of symptoms of only 2.2 years. The ACTIVE trial³⁵ did not report durations. The main evidence comes from the TIG/ACT⁴³ 5-year data for which only those with duration of symptoms under 3 years showed a significant difference between ACI and MF. Improvements in KOOS at 5 years were 26 for the ACI group versus 15 for the MF group ($p = 0.026$). For the subgroup with over 3 years' duration, KOOS improvements were 13 for ACI and 17 for MF (not significant). This might suggest that ACI is of less value, relative to MF, in patients with longer duration.

Previous studies have shown improvements with ACI after long duration of symptoms. In the trial by Bentley *et al.*,⁷⁸ most patients receiving ACI had excellent Cincinnati scores results despite a mean duration of symptoms of 7.2 years. In the trial of ACI-C versus MACI by Bartlett *et al.*,⁵⁷ 59% of the ACI-C group and 72% of the MACI group had good or excellent Cincinnati scores despite duration of symptoms of approximately 10 and 7 years, respectively. In another study from Stanmore by Biant *et al.*,⁷⁹ of a cohort of 104 ACI patients followed for at least 10 years, 66% had excellent or good Cincinnati scores despite an average duration of symptoms before ACI of 7.8 years.

ACI-C or MACI?

In a very large cohort of 827 patients with mean duration of follow-up 6.2 years, Nawaz *et al.*⁸⁰ reported better results with MACI than ACI-P or ACI-C, though this was probably due to different durations, as the ACI groups came from an earlier period and so had more time for knee status to decline. The RCT of ACI-C versus MACI by Bartlett *et al.*⁵⁷ found no difference at 1 year.

In practice, ACI has evolved and most use is now expected to be MACI, with characterised chondrocytes.

Predicting success

Nawaz *et al.*⁸⁰ noted that the chance of success was reduced by previous attempts at repair and the presence of any osteoarthritic change in the knee. They summarised the results from their very large cohort study thus:

Our study suggests that the 'ideal' candidate for autologous chondrocyte implantation is a younger individual with a single lesion on the trochlea or the lateral femoral condyle, with no previous procedures or evidence of degenerative changes.

At 12 years after ACI, their ideal patients had a repair survival of almost 80% compared with 50% in the whole group. Graft survival was < 25% in those who had had a previous repair (including MF and mosaicplasty) and > 75% in those who had not.

Survival of repairs

The 2-year differences between MF and ACI or MACI arise mainly because symptom scores reach a plateau sooner after MF than after ACI. Saris *et al.*⁷⁰ reported (from graph) that a KOOS plateau was reached with MF by 12–18 months, whereas improvements continued after ChondroCelect ACI-P. The SUMMIT¹⁰⁰ investigators showed a plateau before 12 months with MF but not until 18 months with MACI.

In the TIG/ACT trial,⁴³ Saris *et al.*⁷⁰ reported (from graph, so approximate) that about 7% of MF repairs had failed by 20 months and 11.5% by 36 months (but based on only seven failures in the MF group). The longer-term results reported by Vanlauwe *et al.*⁴³ showed the plateau in the KOOS in the MF group from 12 to 60 months, whereas the ChondroCelect group with duration of symptoms < 3 years at surgery, reached a plateau at 36 months. The CC group with duration of symptoms of > 3 years showed no difference in KOOS from MF with an early plateau and lines almost overlapping.

Basad *et al.*⁶⁴ reported that the Lysholm score in the MF group improved from 55 at baseline to 81 at 12 months, but then declined to 69 at 24 months. The MACI group had a baseline score of 55, improving to 95 at 12 months, maintained at 92 at 24 months.

Bhosale *et al.*,¹⁰⁸ from Oswestry, report results at an average of 5 years (range about 3–9 years) among 80 patients, all but four having had ACI-P. The median baseline Lysholm score was 54, which improved to a median of 78 at 12 months post operation. Of the 80 patients, 65 improved and scores at 15 months were maintained for up to 9 years. They also reported that higher age, female gender and larger defect size were associated with greater benefit, but none of these associations was statistically significant. They concluded that a good result at 15 months is durable.

Commentary on SoBi submission October 2015: survival analysis

SoBi pooled reconstructed IPD for five ACI studies that provided data beyond 5 years. The studies varied in failure definition and proportion of patients previously treated. Parametric models were fitted and according to information criteria, the best fit was from a Gompertz model followed by a gamma model. These studies encompassed 507 patients, and included one study with 62 participants, which was not included by the Assessment Group. This study was by Filardo *et al.*,¹⁵⁶ who were using the Hyalograft scaffold, which was a bioengineered non-collagen product that was not covered by the NICE scope. Hyalograft was withdrawn from the market in January 2013. SoBi excluded the largest relevant study (Nawaz 2014⁸⁰) with 827 patients, which we think is the most relevant study because it was undertaken with UK patients; had a mix of ACI generations and patients; and provided subgroup analyses. Also excluded were arms of studies with data to 5 years. When the six ACI studies (which we identified, with 5 years' or more follow-up) are pooled the best fits are provided by a gamma model (*Table 20*).

Figure 16 shows the KM plot and best fit gamma model (95% CI) to 70 years post intervention for the pooled six ACI studies (Knutsen 2007,⁶⁷ Vanlauwe 2011,⁴³ Nawaz 2014,⁸⁰ Niemeyer 2014,¹¹⁷ Minas 2014,¹³⁶ Moseley 2010¹¹⁵), together with the SoBi best- and worst-scenario models based on five ACI studies.

TABLE 20 Information criteria for six studies with ACI arms examined in the Assessment Group report

| Model | Obs. | ll(model) | df | AIC | BIC |
|--------------|------|-----------|----|----------|----------|
| Gamma | 1270 | -975.825 | 3 | 1957.649 | 1973.09 |
| Exponential | 1270 | -1013.96 | 1 | 2029.922 | 2035.069 |
| Weibull | 1270 | -1006.09 | 2 | 2016.178 | 2026.472 |
| Gompertz | 1270 | -1013.86 | 2 | 2031.712 | 2042.005 |
| Log-normal | 1270 | -982.15 | 2 | 1968.301 | 1978.594 |
| Log-logistic | 1270 | -993.514 | 2 | 1991.028 | 2001.321 |

df, degrees of freedom.

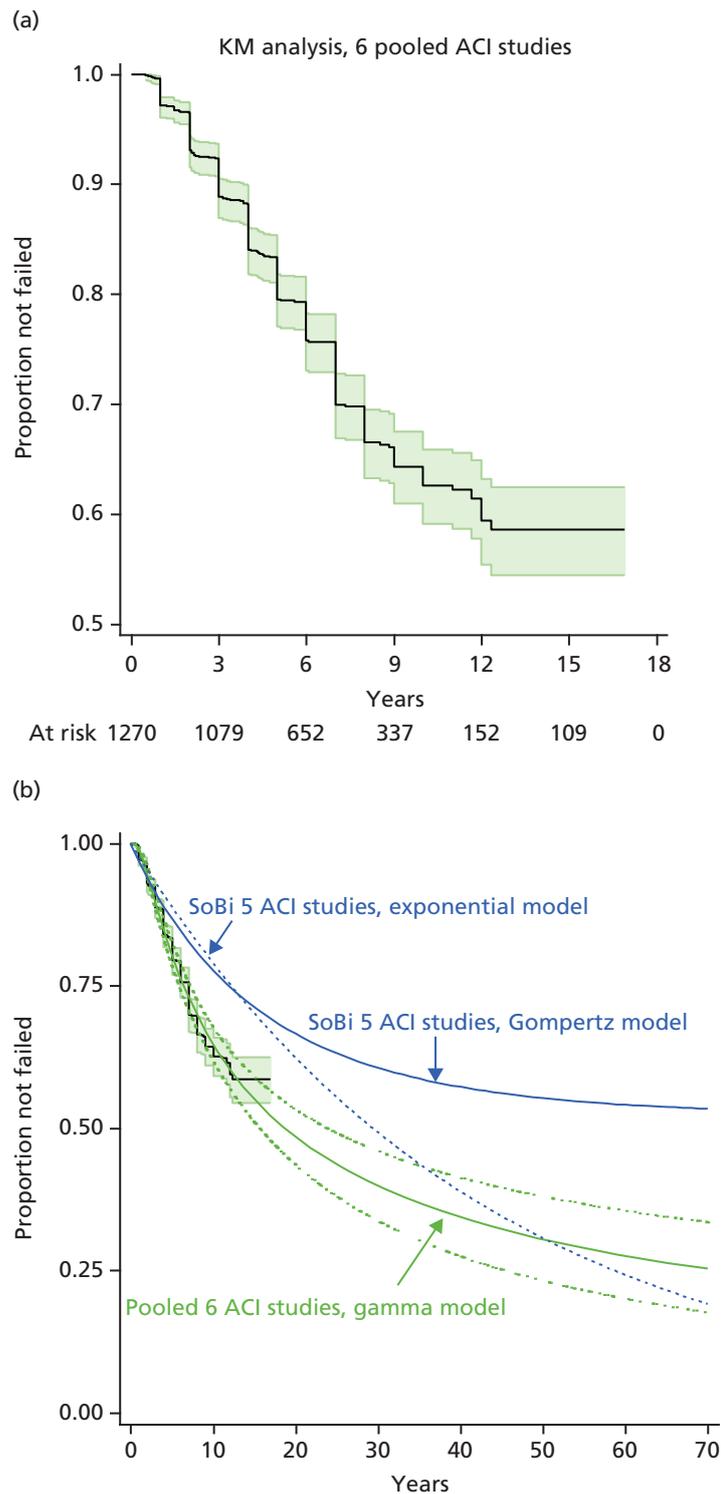


FIGURE 16 Pooled ACI studies compared with SoBi curves. The gamma model for six studies generates poorer survival than the Gompertz model generated by SoBi for five pooled studies.

Alternative models for the six pooled studies are shown in *Figure 17*. It may be a moot question whether pooling the five ACI studies of SoBi or all of the studies that we identified is preferable, but omission of the largest UK study by SoBi does not seem appropriate.

For the purposes of base-case economic modelling, SoBi used the Gompertz fits to five pooled studies to develop a model of failure of ACI for times beyond the 71 months of observed data (KM) from the Vanlauwe TIG/ACT study⁴³ (*Figure 18*). The TIG/ACT study⁴³ included only 51 patients in ACI arm and it

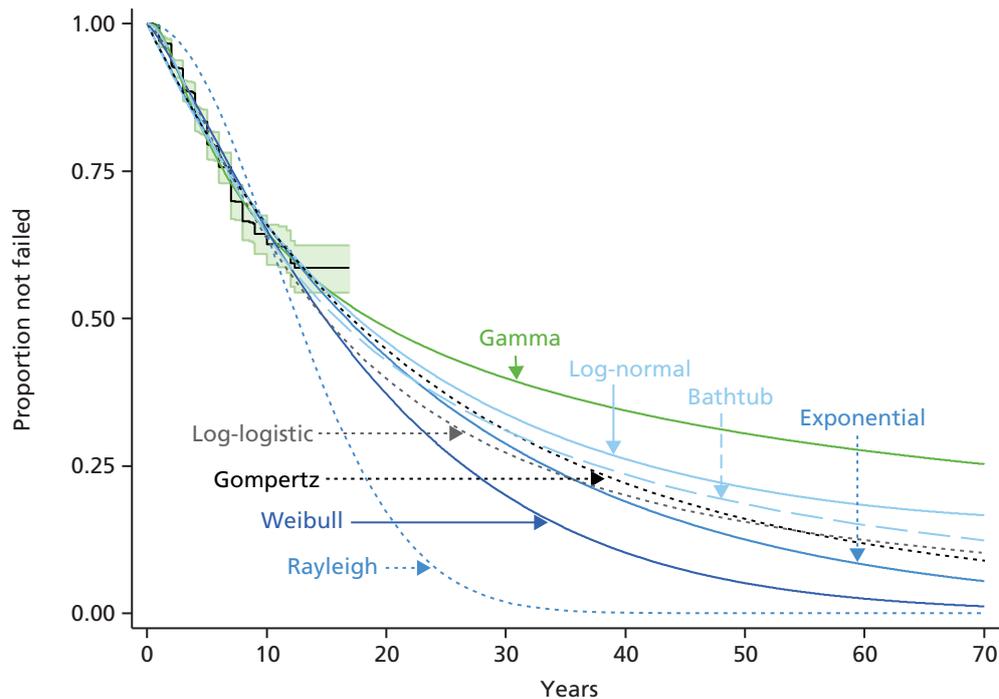


FIGURE 17 Alternative models for six pooled ACI studies ($n = 1270$).

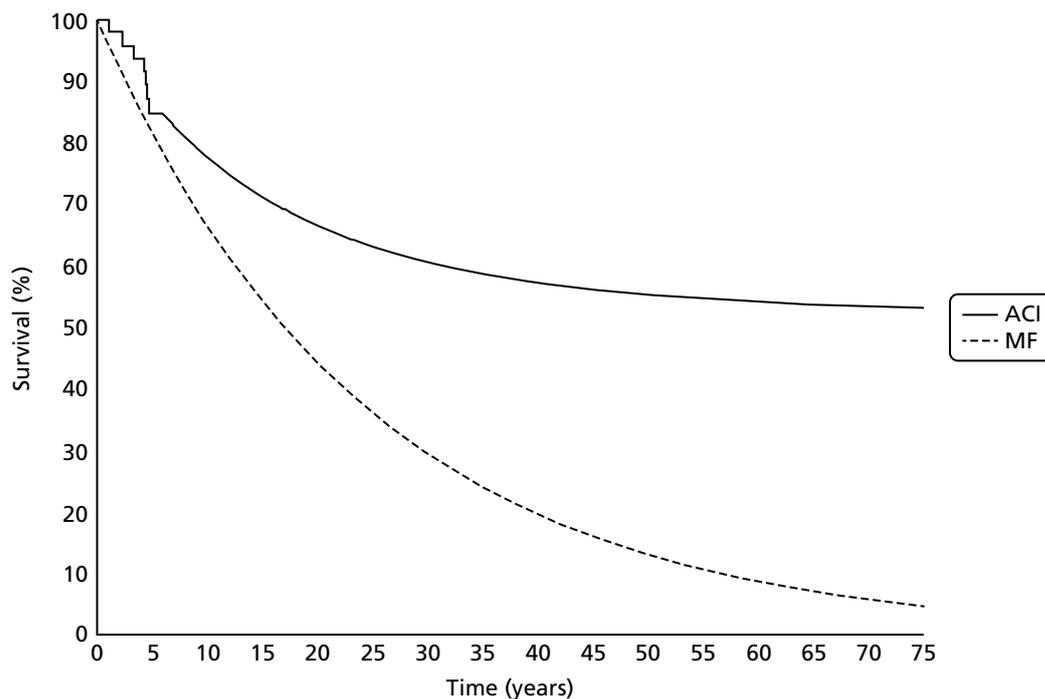


FIGURE 18 SoBi hybrid curve for ACI failure. (TIG/ACT⁴³ KM to 70 months then a Gompertz model for five pooled ACI studies.) Also shown is the SoBi MF failure model based on an exponential fit to the MF arm of the TIG/ACT study.

might be suggested that using the pooled data for all the ACI arms would be more appropriate. The resulting hybrid curve generated by SoBi incorporates data for 51 patients to 71 months and an extrapolation based on a Gompertz curve that excluded these 51 patients. The resulting hybrid may be considered to probably flatter ACI, in that the major Nawaz study⁸⁰ has been excluded.

SoBi have not pooled MF studies. For the comparator MF arm, the new submission appears to have used an exponential survival curve based on the MF arm of the TIG/ACT study⁴³ as in a previous submission; however, this is unclear.

The Assessment Group found an anomaly in the published risk table for the MF arm of the TIG/ACT study.⁴³ A speculative correction to the risk table allowed reconstruction of IPD that yielded the exponential model shown in *Figure 19a*. This plot is closely similar to that proposed by SoBi. Alternative candidate models (*Figure 19b*) produce variously different models of failure. The small number of patients and apparent anomaly in published data render these curves problematical.

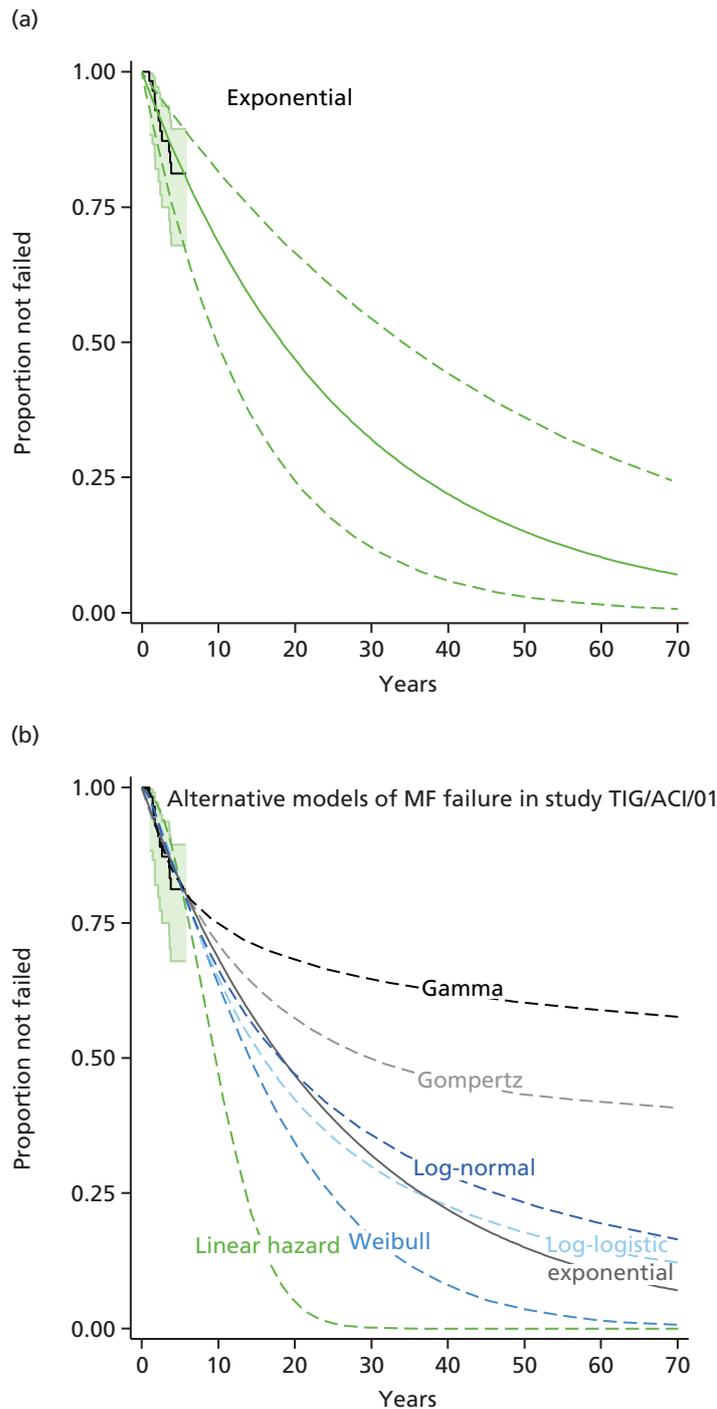


FIGURE 19 Exponential and other parametric models of MF failure based on the MF arm of the TIG/ACT⁴³ study.

Based on their pooled studies, the SoBi analysis concludes that survival of ACI was 70% after 15 years. Applying their long-term survival data in the modelling reduces the ICER from about £26,000 to £21,000, and adding a confidential Patient Access Scheme reduces further.

The weakness in the SoBi analysis is the lack of similar survival analysis for MF, but that is partly because there are fewer long-term studies of MF than of ACI.

Chapter 6 The cost-effectiveness of autologous chondrocyte implantation

Introduction

The first aim of this analysis is to determine whether or not ACI is cost-effective compared with the current standard treatment of MF as primary treatment for patients with symptomatic articular cartilage defects of the knee. We use ACI as a generic term to cover all of the relevant forms of ACI.

After the first procedure, patients may have a number of outcomes:

- Permanent success, more likely with ACI than MF.
- Temporary success followed by a second attempt at repair, or at a longer interval, knee replacement.
- Failure followed by another repair.
- Failure, but the patient may decide against another repair and treat symptoms with analgesics, perhaps because they got some relief from the first repair. He/she would probably develop OA, and might have a knee replacement in later life, ideally not until over 55 years old.

Second repairs could be ACI or MF.

A simplified diagram of the repair options is shown in *Figure 20*. The simplifications are twofold. First, 'success' may not be permanent, especially in the case of MF. Second, this figure does not show longer-term sequelae, such as OA and need for knee replacement. This is shown in the detailed model diagram below (see *Figure 21*). We distinguish repairs, ACI and MF, from replacements such as partial or total knee arthroplasties (PKR and TKR).

Scenario 1 (top) shows that all second repairs are ACI and scenario 2 (bottom) shows that all second repairs are MF. This is to allow a direct comparison between ACI and MF as first procedure. In practice, if a second repair is needed, the choice may vary according to what the first repair was – we deal with other possible sequences later.

This chapter describes the structure of the model, the parameters used within the model (transition probabilities, resource use, costs and utilities), the assumptions made, the different scenarios that have been evaluated, the base-case results and the SAs undertaken.

Model structure

A Markov (state-transition) model was developed in Microsoft Excel® (Microsoft Corporation, Redmond, WA, USA). A Markov model was considered to be the most appropriate, as we wanted to determine whether ACI would postpone or avoid knee replacement in the longer term. The economic model reflects the different clinical pathways for patients with symptomatic articular cartilage defects of the knee. We have used information from the systematic review of cost-effectiveness studies for ACI, most notably Clar *et al.*³ and Gerlier *et al.*,¹⁴¹ and this has been supplemented by information from expert clinical opinion in order to develop the clinical pathways.

In practice, some patients who would be considered for ACI should that be approved will have had a previous procedure, most often MF, but this is covered in the set of sequences below. For those who do

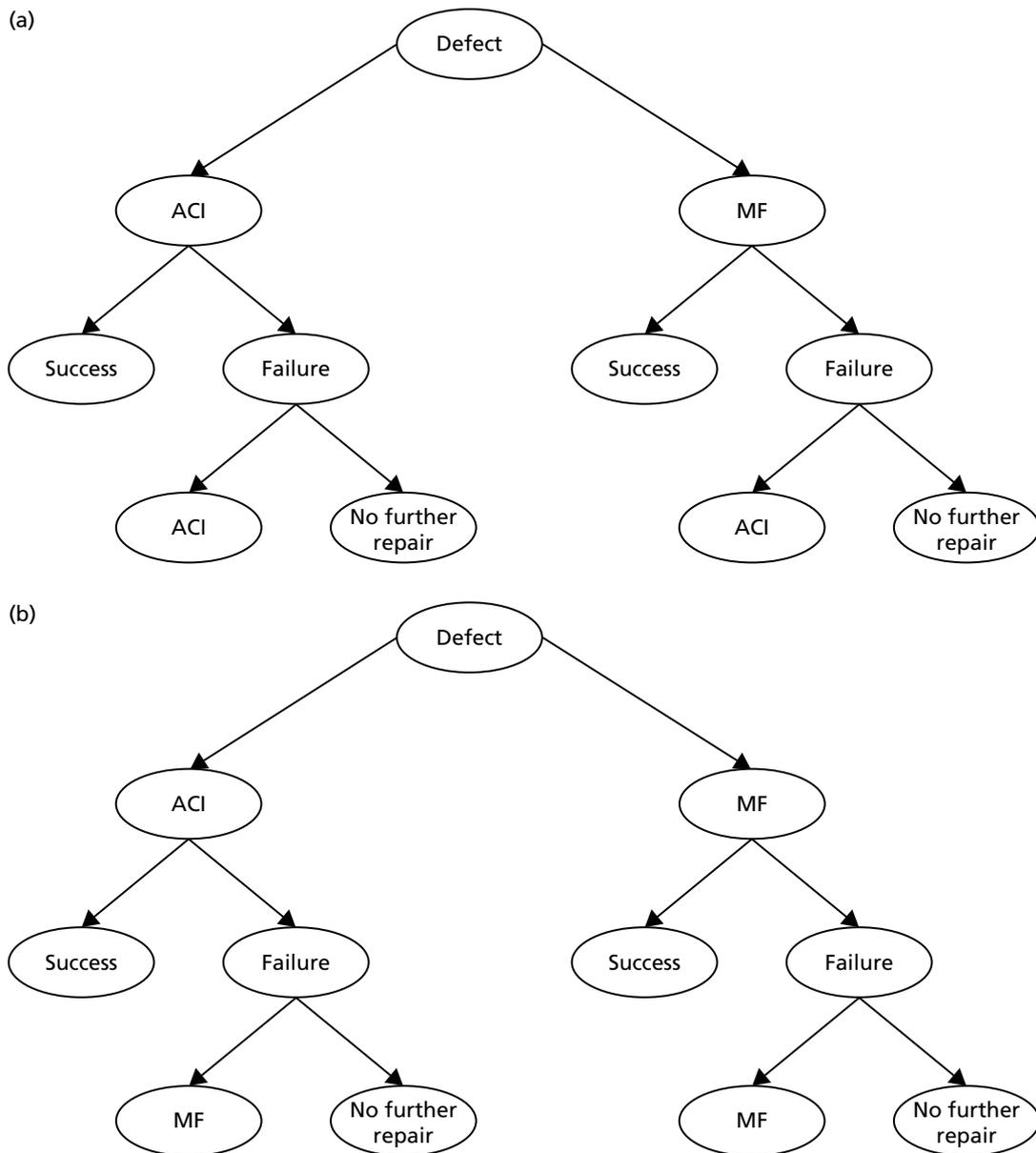


FIGURE 20 Patient pathways for ACI or MF: scenario 1 (above), scenario 2 (below).

need a second repair, we considered both ACI and MF in the sequences within the model. We have assumed that patients will have a maximum of two repairs and combinations could be as follows:

1. ACI(ACI) Patients receive ACI as a primary repair and *if* they require a second repair then this will also be an ACI.
2. MF(MF) Patients receive MF as a primary repair and *if* they require a second repair then this will also be MF.
3. ACI(MF) Patients receive ACI as a primary repair and *if* they require a second repair then this will be MF.
4. MF(ACI) Patients receive MF as a primary repair and *if* they require a second repair then this will be an ACI.

Clinical pathways

Figure 21 shows the detailed clinical pathway for people receiving treatment for symptomatic articular cartilage defects of the knee.

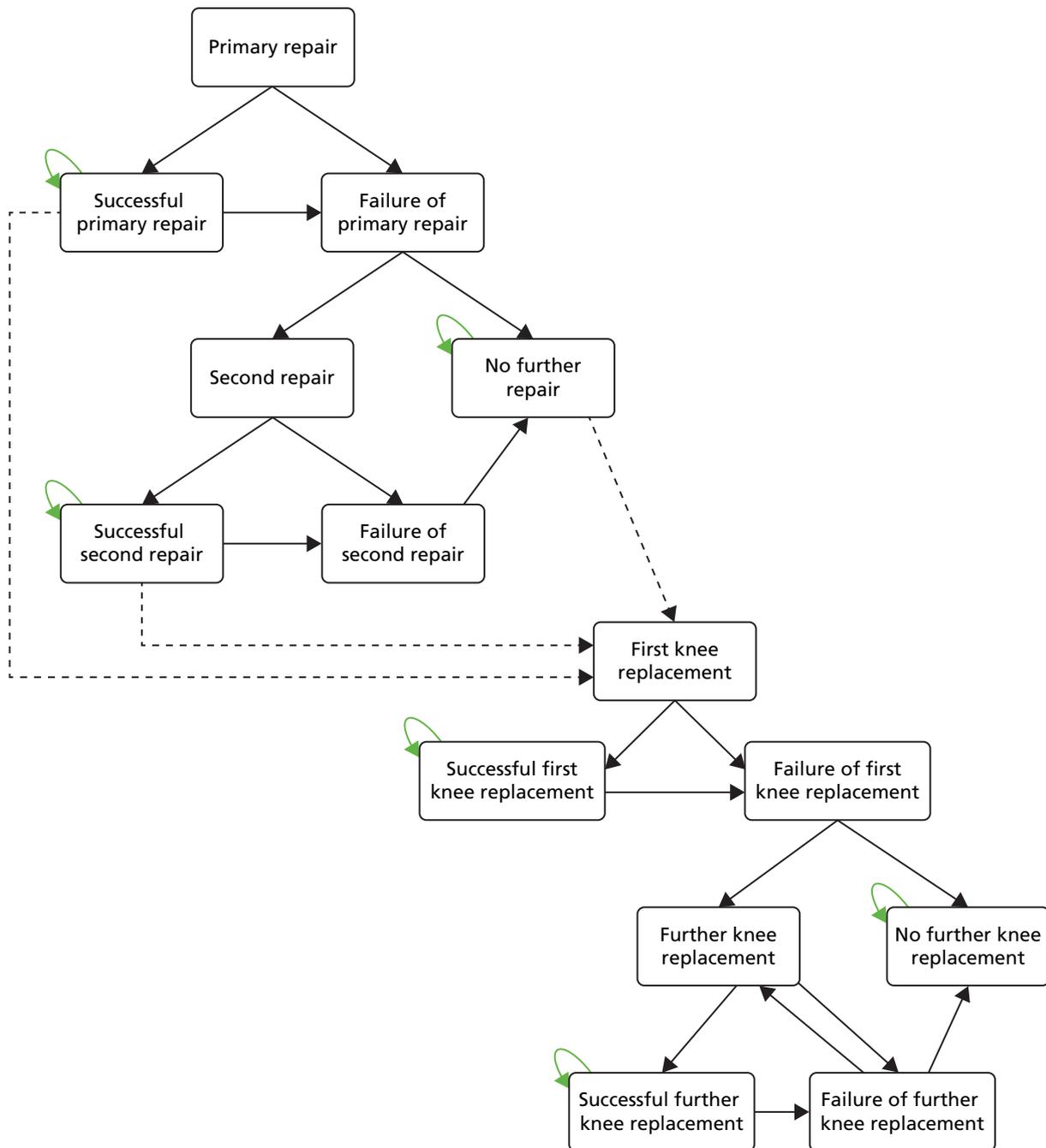


FIGURE 21 Clinical pathways for patients with articular cartilage defects of the knee joint.

- All health states lead to the 'dead' health state.
- The little loop arrow in the top left-hand corner means that this can be temporary or permanent success.
- The dashed line indicates that at 55 years of age, the patient can have a knee replacement (total or partial).

Knee repairs

The starting point for the model is the primary repair, which could be either ACI or MF. After the primary repairs, patients can then either move to the 'successful primary repair' health state or to the 'failure of primary repair' health state. Success can be permanent: the first repair works and they do not require a second repair. So they can stay in the 'successful primary repair' health state until they die. Or success can be temporary (the patient has no symptoms for years but after a while the repair fails), so the patient then moves to the 'failure of primary repair' health state. They can then have a second repair or they can choose not to have another repair (no further repair health state) and rely on analgesics to relieve

symptoms, that is, the patient chooses to accept the pain and treat it rather than have another attempt at repair, though later he/she may have a knee replacement.

The second repair could be either ACI or MF. Based on clinical opinion, we have assumed that patients will have a maximum of two repairs. Once the patient has had a second repair he/she can then either move to the 'successful second repair' health state or to the 'failure of second repair' health state. The successful second repair can be permanent (similar to the successful primary repair) and patients stay in this health state until they die. Or it could be a temporary success, so then the patient later moves to the 'failure of second repair' health state. We are assuming that patients whose second repair fails do not have another repair and they move to the 'no further repair' health state.

Patients who move to the 'no further repair' health state after failure of repair can choose not to have another repair procedure and accept the pain, taking analgesics as required (that is, they can stay in this health state), until they reach the knee replacement age range, when their options are knee replacement or continued symptomatic treatment. Those who choose not to have a further repair may have had partial relief from symptoms, so we rate their utility as better than the baseline one.

Knee replacements

We assume for simplicity that patients over the age of 55 years cannot have an ACI or MF, but have only a knee replacement or symptomatic care. This is in line with the MSAC report, which indicated that MACI/ACI was not indicated for patients older than 55 years.³² The first knee replacement can be either a partial (unicompartmental) knee replacement (PKR) or TKR. According to statistics from the National Joint Registry, the average ages of patients having a PKR and TKR are 64 and 70 years of age, respectively.¹⁵⁷ However, we know that patients being considered for ACI are a lot younger than the general population (average age early thirties), so that if the repair fails, they are more likely to have a knee replacement at an earlier age. In line with expert clinical advice, we are assuming that patients can have one or more knee replacements. The first may succeed for life; if not, they can have another replacement or choose not to have another replacement. The first knee replacement could be either a PKR or a TKR, but we have assumed that all subsequent replacements will be TKRs.

A patient can move to 'first knee replacement' from either a temporarily 'successful primary repair' health state, a temporarily 'successful second repair' health state or from the 'no further repair' health state when they reach the knee replacement age range. The first knee replacement can be a success, so the patient moves to the 'successful first knee replacement' health state or the replacement can fail over time, so they move to the 'failure of first knee replacement' health state. The first knee replacement can be a permanent success until the patient dies or a temporary success because the knee replacement fails over time, so he/she moves to the 'failure of first knee replacement' health state, from which patients can choose to have another knee replacement or to have no further knee replacement (so move to the 'no further knee replacement' health state).

The second knee replacement can be a permanent success (until death) or it could be a temporary success, and patients move to the 'failure of further knee replacement' health state, from which they can choose to have no further knee replacement (but use symptomatic treatment) and or to have another (third) knee replacement. Based on clinical opinion, we have assumed that patients can have more than two knee replacements. Patients who move to the 'no further knee replacement' health state choose not to have another knee replacement and stay in this health state until they die.

Deaths

Patients can move to death from any of the repair and replacement health states due to all-cause mortality, or because of the rare mortality associated with PKR or TKR (such as DVT and pulmonary embolism). The latter becomes more relevant in later stages because replacing previous knee replacements requires more extensive procedures.

Markov model structure

The Markov model structure is shown in *Figure 22*. In line with the clinical pathway shown in *Figure 21*, the model shows the different health states and events that can take place. The different events health states for the model are shown by the ovals. The model shows all of the transitions that can happen between the different health states by the direction of the arrows. The little loop arrows in the left-hand corner of the ovals (recurring arrow) means that a patient can stay in that health state for more than one cycle, and perhaps indefinitely (until they die). The dashed line indicates that at 55 years of age, the patient can choose to have a knee replacement (total or partial). Transition probabilities, that is, the rate of progression from one health state to another (or for staying in the same health state), were identified from the literature.

Base-case analysis

Many people with cartilage injury are young and involved in sports, and this is where most of the injuries occur. We have not differentiated by gender, as evidence shows that there is no difference in the success or failure of the two different procedures (ACI and MF) if lesions are comparable.¹⁴⁶ For the base-case analysis, we have adopted a lifetime horizon (i.e. patients can live to 100 years), with a cycle length for the model set at 1 year, and transitions between each health state occurring at the end of each cycle. A cycle length of 1 year is reasonable, given the time that it takes patients to recover from surgery. A hypothetical cohort of a 1000 patients with symptomatic articular cartilage defects of the knee with a starting age of 33 years is followed from their first repair. The analysis is conducted from the perspective of the NHS and PSS. All costs are in UK pounds sterling in 2012–13 prices. Health outcomes were measured in QALYs. Results are expressed as incremental cost per QALY gained. An annual discount rate of 3.5% is applied to both costs and outcomes.

Model inputs

Transition probabilities

For the base-case analysis, annual transition probabilities were based on data derived from the literature and in consultation with clinical experts. For the primary and second repairs for both ACI and MF, these transition probabilities were based on success rates for ACI compared with MF, and these probabilities came from two main studies: Saris *et al.* 2009⁷⁰ and 2014.¹⁰⁰

Figure 23 shows a flow chart with the proportion of people achieving success or failure with each repair. Saris *et al.*⁷⁰ reported a success rate of 83.0% over a 3-year period for ACI, and, of the remaining 17% of people, 3.9% would require reoperation of the same lesion (either ACI or MF as a second repair). We therefore assumed that 13.1% of people would have no further repair following the first repair. After ACI as a first repair, for those patients who required a second repair, and if this repair was another ACI, we assumed the same success rate for this second repair as the first repair. If this second repair was MF, we used information from Vanlauwe *et al.*,⁴³ who reported that for 16.4% of people MF after a first ACI failed at 5 years. We assumed that the rest of the people had successful MF procedures following an ACI.

Saris *et al.*⁷⁰ also reported a success rate of 62.0% over a 3-year period for MF and, of the remaining 38% of people, 11.5% would require reoperation of the same lesion (either ACI or MF as a second repair). We therefore assumed that 26.5% of people would have no further repair following the first repair. After MF as a first repair, for those patients who required a second repair and if this repair was MF, we assumed the same success rate for this second repair as the first repair. If this second repair was an ACI, this information was obtained from Biant *et al.*,⁷⁹ who reported that for 30.9% of people an ACI after a first MF failed at 10 years. We assumed that the rest of the people had a successful ACI procedure following MF.

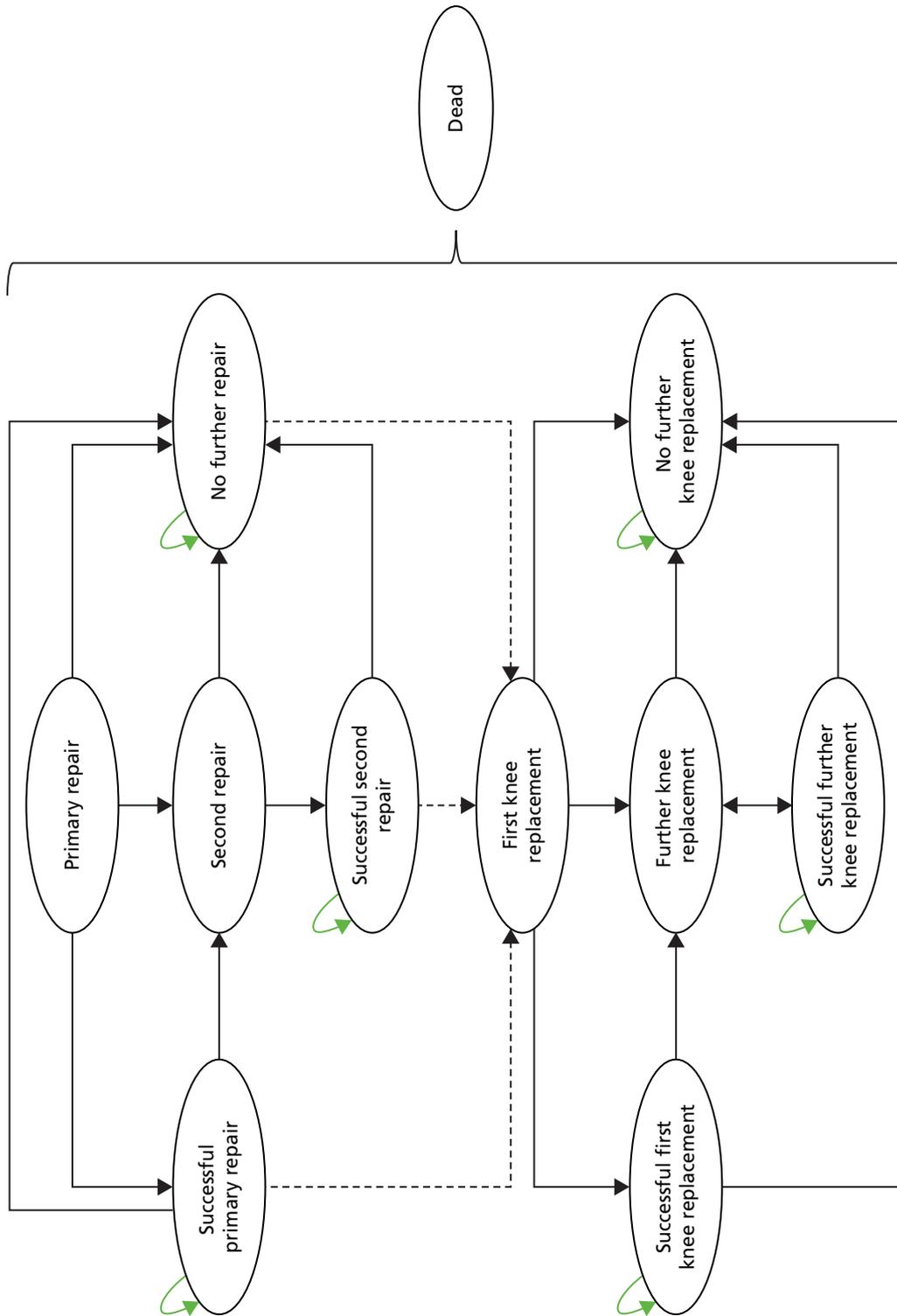


FIGURE 22 Markov model structure for patients with articular cartilage defects of the knee joint.

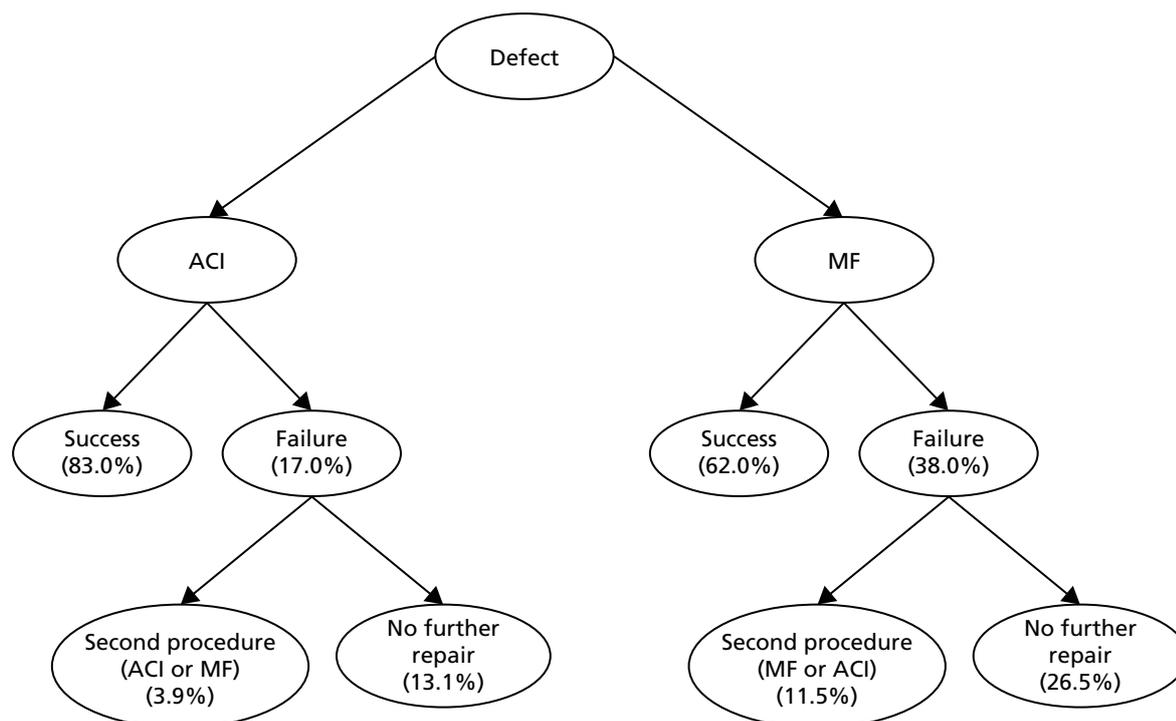


FIGURE 23 Proportion of patients achieving success/failure with ACI or MF at 36 months.

The timing of knee replacement after one of the repair health states was based on data from the RCT of ACI and MF by Knutsen *et al.*⁶⁷ Transition probabilities for success and failure for patients who needed knee replacements or knee replacement revisions were derived from two studies: Gerlier *et al.*¹⁴¹ and Dong and Buxton.¹⁵⁸ *Appendix 11* details the literature used and the assumptions made for deriving these probabilities, and *Tables 57–59* show the transition probabilities that have been used in the base-case analysis.

Utilities

There are very few studies reporting health-state utility values for patients with symptomatic articular cartilage defects of the knee. The main studies reporting utility values have been summarised in *Chapter 4* (Clar 2005,³ Derrett 2005,¹⁴⁰ Gerlier 2010¹⁴¹). In the previous HTA report, the preoperative QoL value was taken to be 0.80 and for those who had successful knee repair there was a utility gain of 0.10 (utility score for successful knee repair was 0.90); for those where the knee repair failed the utility value remained at the preoperative value (utility score for knee repair failure was 0.80) (Clar 2005³). Derrett *et al.*¹⁴⁰ used the EQ-5D-3L to elicit utility scores, and the ACI waiting list group had a pre-surgery utility score of 0.41. After surgery, the EQ-5D-3L mean score for the ACI group was 0.64 and for mosaicplasty was 0.47, a utility gain of 0.23 and 0.06, respectively.

For our model we have used utility values for knee repairs from the Gerlier *et al.* study,¹⁴¹ comparing ACI with MF using data from the TIG/ACT⁴³ ChondroCelect trial. They used a short-term model with a time horizon of 5 years to take into account knee pain and mobility after the initial intervention (QoL information was obtained from a 5-year RCT using the SF-36) and also a long-term model with a time horizon of 40 years to take into account the development of OA after 15 years and the need for a TKR after 20 years. We used two other studies to supplement utility values for knee replacement. The first study is by Dong and Buxton,¹⁵⁸ who developed a Markov model to compare the cost-effectiveness of TKR using computer-assisted surgery with that of TKR using a conventional manual method in the absence of formal clinical trial evidence. The second study is by Jansson and Granath,¹⁵⁹ who analysed EQ-5D data before and after knee arthroplasty.

Table 21 shows the base-case mean utility values used in the model. For the repairs these values were all obtained from the paper by Gerlier *et al.*,¹⁴¹ who used the SF-36 and the KOOS measures to estimate utility scores. The mean utility value for patients before they have a primary repair (before ACI or MF) was 0.654: this utility value was based on the initial value before the intervention. For those patients who had an ACI as a first repair and moved to the 'successful primary repair' health state, we assumed that the patients' mean utility value after surgery for the first year would be 0.760 (this value was based on year 1 post intervention, regardless of the outcome, and takes into account the long rehabilitation period and abstinence from active pursuits), and, if they remain in this health state in subsequent years, their utility value would remain constant at 0.817. This latter value was based on patients who had clinical success for

TABLE 21 Base-case mean utility values used in the economic model

| Repairs/replacements | | First repair | | | | Source |
|---|-----------------|--------------------|-------|-------|-------|---|
| | | ACI | MF | | | |
| Repairs | | | | | | |
| Before primary repair | | 0.654 | | | | Gerlier 2010 ¹⁴¹ |
| Successful primary | First year | 0.760 | 0.760 | | | Gerlier 2010 ¹⁴¹ |
| | Second year | 0.817 | 0.817 | | | |
| | Third year | 0.817 | 0.817 | | | |
| | Fourth year | 0.817 | 0.817 | | | |
| | Five years plus | 0.817 | 0.654 | | | |
| Before second repair | | 0.654 | | | | |
| Choose not to have a second repair | | 0.691 ^a | | | | |
| | | Second repair | | | | |
| | | ACI | MF | ACI | MF | |
| Successful second | First year | 0.760 | 0.760 | 0.760 | 0.760 | Gerlier 2010 ¹⁴¹ |
| | Second year | 0.817 | 0.817 | 0.817 | 0.817 | |
| | Third year | 0.817 | 0.817 | 0.817 | 0.817 | |
| | Fourth year | 0.817 | 0.817 | 0.789 | 0.817 | |
| | Five years plus | 0.817 | 0.654 | 0.789 | 0.654 | |
| No further repair | | 0.691 | | | | |
| Replacements | | | | | | |
| Before first knee replacement (TKR) | | 0.615 | | | | Dong and Buxton 2006, ¹⁵⁸ Jansson and Granath 2011 ¹⁵⁹ |
| Before first knee replacement (PKR) | | 0.615 | | | | Dong and Buxton 2006, ¹⁵⁸ Jansson and Granath 2011 ¹⁵⁹ |
| Successful first knee replacement – TKR | | 0.780 | | | | Dong and Buxton 2006 ¹⁵⁸ |
| Successful first knee replacement – PKR | | 0.780 | | | | Dong and Buxton 2006 ¹⁵⁸ |
| Before further TKR | | 0.557 | | | | Gerlier 2010 ¹⁴¹ |
| Successful further TKR | | 0.780 | | | | Dong and Buxton 2006 ¹⁵⁸ |
| No further TKR | | 0.691 | | | | Gerlier 2010 ¹⁴¹ |

^a Some patients decide not to have another repair attempt after unsuccessful first repair. We have assumed that they had some benefit and do not go back to their baseline utility.

5 years after the intervention. For those patients who had MF as a first repair and moved to the 'successful primary repair' health state, we assumed that the patients' mean utility value after surgery for the first year would be 0.760 (this value was based on year 1 post intervention, regardless of the outcome). For years 2–4 after MF this mean utility value would increase to 0.817. This reflects the quite long rehabilitation required in the first year after the procedures, and the time taken for the cartilage to be replaced. For years 5 and onwards for patients who stay in this same health state, we have assumed that utility would fall to the same as that pre-surgery (mean utility value is 0.654) because the benefit of MF declines after 5 years and patients may choose to have another procedure.

For those patients who require a second repair, the mean utility value was 0.654: this value was based on the utility value before the intervention.¹⁴¹ For those requiring a second repair there are four possible sequences: ACI(ACI), ACI(MF), MF(ACI) and MF(MF).

Utilities for patients having a second successful ACI after the first ACI were assumed to be the same as for those who had a successful ACI as a first repair. Utilities for patients having a successful MF after an initial ACI that failed were assumed to be the same as for those who had a successful MF as a first repair and moved to the 'successful primary repair' health state.

Patients who have a successful ACI after an initial MF move to the 'successful second repair' health state. However, as noted previously, ACI is less effective in patients who have had prior MF, so for years 4 and 5 we have used the average of two utility values from Gerlier *et al.*,¹⁴¹ based on year 1 post intervention (utility value = 0.760) and clinical success after 5 years following the intervention (utility value = 0.817), so the mean utility value for ACI after MF was 0.789. Utilities for patients having a successful second MF after a failed initial MF and who moved to the 'successful second repair' health state were assumed to be the same as those who had a successful MF as a first repair.

For patients who moved to the 'no further repair' health state, the mean utility value was 0.691: this value was based on patients who had not had a successful result 5 years after surgery (Gerlier 2010¹⁴¹), but we have assumed that those who choose to have no further repair may have had some benefit from the first repair, and so do not go back as far as the original baseline utility.

Mean utility values are the same for knee replacements after ACI or MF. Before the first knee replacement procedure, patients who received a TKR and PKR are assumed to have the same utility value, 0.615. This value was based on an average of two utility values: (1) the EQ-5D index score at baseline preoperatively for knee arthroplasty (value = 0.51)¹⁵⁹ and (2) an estimated value for TKR operation for knee problem (value = 0.72).¹⁵⁸ For patients who move to the 'successful first knee replacement' health state (TKR or PKR), a utility value of 0.780 was also obtained from Dong and Buxton.¹⁵⁸ This utility value was estimated from the generic Knee Society Score scale and was applied to the Markov health state for normal health after primary TKR. We have also assumed that if patients move to the successful 'further total knee replacement' health state then they will have the same utility value as if it was a first TKR. For those patients for whom TKR has failed and need a further TKR, the utility value was 0.557 based on the failed TKR/revision health state from Gerlier *et al.*¹⁴¹ Finally, for those patients who move to the 'no further replacement health state' this value (mean = 0.691) was also from Gerlier *et al.*¹⁴¹ and was based on patients who had no clinical success 5 years after surgery (in line with patients who move to the 'no further repair' health state).

Resource use and costs

Costs for the different procedures (ACI, MF, PKR/TKR, TKR revisions) and for outpatient visits and rehabilitation are shown in *Table 22*. We have used national reference costs where possible (NHS reference costs 2013¹⁵¹) supplemented by the previous HTA report on ACI³. All unit costs are presented in pounds sterling (£) in 2012–13 prices.

The cost of the ACI (ChondroCelect and MACI) includes the costs associated with cell development, including the ACI kit, staff time and transporting the cells to and from the laboratory. ACI involves two

TABLE 22 Base-case mean costs (UK pounds sterling) used in the economic model

| Procedure | Information | Unit cost (£) | Source |
|------------------------|---|---------------------|------------------------------------|
| ChondroCelect and MACI | Product including courier services and development of cell culture | 16,000 | UK price for ChondroCelect |
| | Procedure 1 – arthroscopy and cell harvest | 710 ^a | Clar 2005 ³ |
| | Procedure 2 – arthrotomy (day case) | 1030 ^a | |
| | <i>Total cost</i> | 17,740 | |
| MF | Procedure (inpatient) | 3020 ^a | Clar 2005 ³ |
| First TKR (PKR or TKR) | HRG code: HB21C – major knee procedures for non-trauma, category 2, without complications | 5676 | NHS reference costs ¹⁵¹ |
| Further TKR | Second TKR | 12,959 ^a | Clar 2005 ³ |
| Outpatient visit | HRG code: WF01A – non-admitted face-to-face consultant led outpatient attendance | 102 | NHS reference costs ¹⁵¹ |
| Rehabilitation | HRG code: REHABL2 – rehabilitation for joint replacement | 256 | |

HCHS, Hospital and Community Health Services; HRG, Healthcare Resource Group.
 a Cost inflated to 2012–13 prices using the HCHS index.¹⁵⁴

procedures: the arthroscopic cell harvest and the re-implantation during arthrotomy. We assumed both would be done as day cases. Based on clinical experience, we have also included the costs of six outpatient visits and three rehabilitation visits in the first year (*Table 23*).

The cost of MF procedure (including an inpatient stay) was obtained from Clar *et al.*³ and the cost has been updated to 2012–13 prices using the Hospital and Community Health Services (HCHS) index.¹⁵⁴ The inpatient stay is required because, unlike after ACI, the patients can have considerable pain after MF because of the drilling into bone. Over the course of the year, the patient would also have three outpatient visits and three rehabilitation visits, and these costs have been added for this health state (based on clinical experience).

The cost for a first knee replacement was obtained from the NHS reference costs¹⁵¹ and we have assumed that it could be either a TKR or a PKR. After a TKR, a subsequent TKR is almost double the cost, because it is technically more difficult. After a PKR, a second knee replacement would be a TKR, and we have assumed that this would cost £5676. If the patient required any more subsequent knee replacements (all of which would be TKRs) then these would cost £12,959. Based on consultation with clinical experts, in the first year after knee replacement, we have included the cost of two outpatient visits (see *Table 23*).

Unit costs were obtained from the NHS reference costs.¹⁵¹

TABLE 23 Base-case resource use for economic model

| Components (over a year) | Procedure | | | Source |
|--------------------------|-----------|----------------|------------------|-------------------------|
| | ACI | MF | TKR | |
| Inpatient days | 0 | 1 ^a | 4.5 ^a | Expert clinical opinion |
| Outpatient visits | 6 | 3 | 2 | |
| Rehabilitation visits | 3 | 3 | 0 | |

a The cost of inpatient stay has been included in the cost for the different procedures.

We have assumed that there will be no further costs after year 1 once patients enter the successful health states ('successful primary repair', 'successful second repair', 'successful first knee replacement' and 'successful further knee replacement'), as patients incur costs such as outpatient or rehabilitation visits during the first year of either a knee repair or a knee replacement. In addition, for the 'no further repair' health state or the 'no further knee replacement' health state, we have not added any costs for the analgesics based on advice from our clinical experts, as these costs are negligible and these patients are not followed up routinely, and it is up to the GP to refer the patient back to the hospital for a knee repair or a knee replacement.

Complications

Adverse events have not been included, as there were no important differences between the two treatment arms.

Mortality

Age-specific mortality rates used in the economic model were based on the UK general population lifetime tables from the ONS.¹⁶⁰ Using the ONS data, the average probabilities of death for men and women were combined. As the cohort ages, mortality rates generally increase throughout the time horizon in the model. In the model, patients from any health state can move to the 'dead' health state. Patients undergoing a knee replacement are subject to a mortality risk during surgery. To reflect this higher mortality, rates were obtained from a study by Mahomed *et al.*¹⁵⁰ For those patients undergoing a TKR and a TKR revision, the mortality rates were reported as 0.7% and 1.1%, respectively.

Measuring cost-effectiveness

The base-case analysis assessed the cost-effectiveness of ACI compared with MF. We calculated for a cohort of patients the expected quality-adjusted survival based on their likelihood of surviving each cycle, their expected health-state utility value, and their expected costs. We have adopted a lifetime horizon from a starting age of 33 years. The analysis is conducted from the perspective of the UK NHS and PSS. Costs are expressed in 2012–13 UK pounds sterling. The main outcome of interest was the QALY. The different sequences of procedures were ranked in order of increasing cost. We eliminated any categories for which another category was cheaper and more effective (simple dominance). If the ICER for a given category was higher than that of the next more effective alternative then this category was eliminated (extended dominance). For the remaining options, we reported the ICERs, measured as cost per QALY gained. Discount rates of 3.5% were applied to both future costs and benefits, as costs and benefits accrued in the future are valued at less than those accrued today.

Sensitivity analysis assesses the uncertainty in parameter inputs used in the Markov model and to check whether or not the results obtained are robust. We present both deterministic and probabilistic results. For the deterministic analysis, we identified the key factors driving the cost-effectiveness. For the PSA, to reflect the amount and pattern of the variation, the analysis attributes probability distributions randomly around specified parameters with simulations, which are repeated to generate ICERs. The PSA was undertaken using 1000 simulations. We used the gamma distribution for costs and the beta distribution for utility values and transition probabilities.¹⁶¹ As the values for costs, utilities and transition probabilities used in the model were means or weighted averages, an assumption was made for the SE in order to calculate the alpha and beta values, which are required for the PSA. For example, we have assumed the SE to be 0.1 of the mean value.^{162,163} These bootstrapped simulations obtained from the PSA were used to construct cost-effectiveness acceptability curves (CEACs) to illustrate the effect of sampling uncertainty in which individual model parameters were sampled from the appropriate probability distribution. CEACs were presented to indicate the probability of a procedure being cost-effective using a willingness-to-pay threshold from £0 to £60,000.

Scenario and sensitivity analyses

Several scenario and sensitivity analyses were conducted by altering base-case inputs to the model.

- SA1 In the base-case analysis, the cost of cells for ChondroCelect and MACI procedures was £16,000. We are aware that confidential discounts are provided to the NHS by manufacturers. So in the SA we have varied this figure by reducing the costs by 25%, 50% and 75%, so that the costs of cells are £12,000, £8000 and £4000, respectively. Note that the cost of cell production in the OsCell laboratory in Oswestry is £4125 per patient.
- SA2 In the base-case analysis, a lifetime horizon was chosen, with the starting age of 33 years for the cohort. In the SA we have varied the time horizon (10, 20, 30, 40 and 50 years) to see how this affects the ICER.
- SA3 In the base-case analysis, according to clinical advice we costed MF as an inpatient procedure (£3020). However, we know that sometimes this procedure is done as a day case. In the SA we have assumed that MF is done as a day case procedure and the associated cost is £1034.
- SA4 In the base-case analysis, the success rates for MF were based on existing evidence. However, there are new types of MF procedures and these could have better success rates. We have no evidence for this, but in a 'what if' SA we have checked what would happen to ICERs if the success rates for MF could increase by 20% and 40%. The effect are to increase duration of benefit after MF. In addition, we assumed that the utility benefit for MF lasted longer as well: for the 20% increase in year 5, and the 40% increase in years 5 and 6, we applied a utility value of 0.817, as opposed to 0.654.
- SA5 In the base-case analysis the starting age for the cohort was 33 years. In the SA the starting age is changed to 45 years (patients are nearer to the knee replacement age) to see how this affects the ICER.
- SA6 In the base-case analysis we used utility values from the paper by Gerlier *et al.*,¹⁴¹ who compared ACI with MF. In this SA we have used utility values that are from the ACTIVE trial³⁵ (Oswestry submission).

Results

We present here the cost-effectiveness deterministic and probabilistic results for ACI compared with MF.

Base-case cost-effectiveness results

One thousand patients entered the model, with a starting age of 33 years. For the primary repair these patients can receive either an ACI or MF, and if these patients require a second repair it could either be an ACI or MF. Many will not require a second repair, but the cost-effectiveness of the primary repair depends partly on the costs of subsequent interventions required or avoided, so we need to consider the sequence options.

Table 24 shows the base-case deterministic and probabilistic cost-effectiveness results for the lifetime horizon for the two different scenarios. For scenario 1, if patients required a second repair this would be an ACI and for scenario 2 if patients required a second repair this would be MF. After MF, 11.9% of patients required a second procedure, and after ACI 3.9% of patients required a second repair. Looking at the discounted deterministic results, for scenario 1, ACI costs £14,314 more than MF, but generated 0.9944 more QALYs than MF. The cost per QALY gained for ACI compared with MF was £14,395. For scenario 2, ACI, again, was more costly (incremental cost = £14,877), but generated more QALYs (0.9537) and the resulting cost per QALY gained was £15,598. For both scenarios, ACI as a first repair was more cost-effective than MF as a first repair. These results were of similar magnitudes and directions for both the undiscounted deterministic results and the probabilistic results.

One of the key cost drivers was the cost of the cells for the ACI procedure, but, over the lifetime horizon, there are QALYs gained from using ACI, and there are cost savings to the NHS later due to fewer people needing a second repair, fewer people in need of a TKR, and fewer people moving to the no further repair/replacement health states (in which the utility is lower).

TABLE 24 Base-case deterministic and probabilistic cost-effectiveness results (by scenario)

| Procedure | Total mean costs (£) | Total mean QALYs | Incremental costs (£) | Incremental QALYs | ICER (£) (cost per QALY gained) |
|------------------------------------|----------------------|------------------|-----------------------|-------------------|---------------------------------|
| Deterministic: undiscounted | | | | | |
| <i>Scenario 1</i> | | | | | |
| MF(ACI) | 8028 | 34.1648 | – | – | – |
| ACI(ACI) | 22,252 | 35.7922 | 14,524 | 1.6273 | 8925 |
| <i>Scenario 2</i> | | | | | |
| MF(MF) | 6234 | 34.1259 | – | – | – |
| ACI(MF) | 21,155 | 35.6504 | 14,921 | 1.5245 | 9788 |
| Deterministic: discounted | | | | | |
| <i>Scenario 1</i> | | | | | |
| MF(ACI) | 6607 | 17.0284 | – | – | – |
| ACI(ACI) | 20,921 | 18.0228 | 14,314 | 0.9944 | 14,395 |
| <i>Scenario 2</i> | | | | | |
| MF(MF) | 5015 | 17.0033 | – | – | – |
| ACI(MF) | 19,892 | 17.9570 | 14,877 | 0.9537 | 15,598 |
| Probabilistic: discounted | | | | | |
| <i>Scenario 1</i> | | | | | |
| MF(ACI) | 6624 | 16.9878 | – | – | – |
| ACI(ACI) | 20,838 | 18.0343 | 14,214 | 1.0466 | 13,581 |
| <i>Scenario 2</i> | | | | | |
| MF(MF) | 5030 | 16.9654 | – | – | – |
| ACI(MF) | 19,809 | 17.9490 | 14,779 | 0.9836 | 15,026 |

Figure 24 presents the CEACs for the base-case results for scenarios 1 and 2, respectively. For scenario 1, if the decision-maker was willing to pay £14,000, the probability that both ACI and MF were cost-effective was approximately 50%; however, if the decision-maker was willing to pay £20,000, ACI was probably 59% more likely to be cost-effective than MF (see Figure 24a). These results were similar for scenario 2: if the decision-maker was willing to pay £20,000, the probability that ACI was more cost-effective than MF was 56% (see Figure 24b).

Table 25 shows the base-case deterministic and probabilistic cost-effectiveness results for the lifetime horizon, ranked by the least costly sequence (option). For the discounted deterministic results MF(MF) was the least costly option and had the fewest QALYs, whereas ACI(ACI) was the most expensive option but generated more QALYs. The option MF(ACI) was extendedly dominated by a linear combination of MF(MF) and ACI(MF) and therefore this option was eliminated from the comparison. The ICER ACI(MF) and MF(MF) was just under £15,000; doing ACI first is more cost-effective. The ICER between the two initial ACI options: ACI(ACI) vs. ACI(MF) was just under £16,000; even if the first ACI fails, there is good enough chance of a second ACI succeeding to make the ICER for a repeat ACI quite reasonable. So initial ACI appears more cost-effective than initial MF and for those who need a second repair after the first ACI, this should be another ACI. These results were of similar magnitudes and directions for the probabilistic results and also for the undiscounted deterministic results.

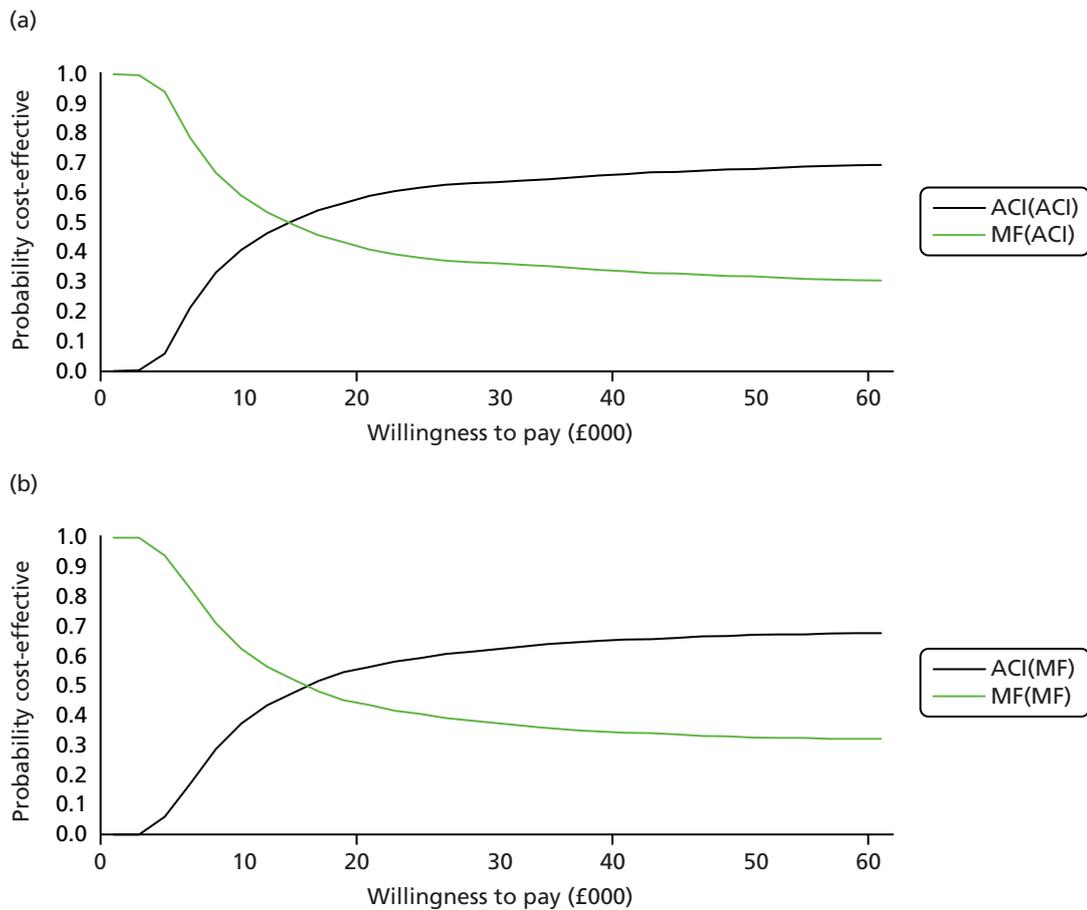


FIGURE 24 Cost-effectiveness acceptability curve: base case results: (a) scenario 1; and (b) scenario 2.

TABLE 25 Base-case deterministic and probabilistic cost-effectiveness results

| Procedure | Total mean costs (£) | Total mean QALYs | Comparison | Incremental costs (£) | Incremental QALYs | ICER (£) (cost per QALY gained) |
|------------------------------------|----------------------|------------------|----------------------|-----------------------|---------------------|---------------------------------|
| Deterministic: undiscounted | | | | | | |
| MF(MF) | 6234 | 34.1259 | – | – | – | – |
| MF(ACI) | 8028 | 34.1648 | MF(ACI) vs. MF(MF) | 1795 | 0.0389 | Extendedly dominated |
| ACI(MF) | 21,155 | 35.6504 | ACI(MF) vs. MF(MF) | 14,921 | 1.5245 | 9,788 |
| ACI(ACI) | 22,252 | 35.7922 | ACI(ACI) vs. ACI(MF) | 1,397 | 0.1418 | 9,856 |
| Deterministic: discounted | | | | | | |
| MF(MF) | 5015 | 17.0033 | – | – | – | – |
| MF(ACI) | 6607 | 17.0284 | MF(ACI) vs. MF(MF) | 1592 | 0.0251 ^a | Extendedly dominated |
| ACI(MF) | 19,892 | 17.9570 | ACI(MF) vs. MF(MF) | 14,877 | 0.9537 | 15,598 |
| ACI(ACI) | 20,921 | 18.0228 | ACI(ACI) vs. ACI(MF) | 1,029 | 0.0658 | 15,648 |
| Probabilistic: discounted | | | | | | |
| MF(MF) | 5030 | 16.9654 | – | – | – | – |
| MF(ACI) | 6624 | 16.9878 | MF(ACI) vs. MF(MF) | 1595 | 0.0223 ^a | Extendedly dominated |
| ACI(MF) | 19,809 | 17.9490 | ACI(MF) vs. MF(MF) | 14,779 | 0.9836 | 15,026 |
| ACI(ACI) | 20,838 | 18.0343 | ACI(ACI) vs. ACI(MF) | 1,029 | 0.0853 | 12,059 |

a As the incremental QALYs are near zero, the ICER can fluctuate widely.

Figure 25 presents the CEAC for the base-case results for all sequences. The graph shows that for amounts below £14,000 then MF(MF) appears cost-effective compared with the other three options. At a willingness to pay of £16,000, there is not much difference between the four options. However, if the decision-maker is willing to pay £18,000 or more for a QALY then ACI as a first procedure [either ACI(ACI) or ACI(MF)] is probably more cost-effective than MF [either MF(ACI) or MF(MF)] as a first procedure.

Scenario and sensitivity analysis cost-effectiveness results

This section highlights the results from the different SAs that were undertaken.

Sensitivity analysis 1: cell cost reduction

In the base-case analysis, the cost of cells for ChondroCelect and MACI procedures was taken to be £16,000. We know that there are confidential discounts provided to the NHS by manufacturers. In this SA we have varied this figure by reducing the cell costs by 25% (£12,000), 50% (£8000) and 75% (£4000). The last figure may seem very low, but it is similar to the cost provided in the Oswestry submission for cells produced in a NHS facility.

Table 26 shows the results when the cost of cells is reduced. When there was a reduction in cell costs, even though ACI was more costly than MF, there were more QALYs gained with ACI than MF. For a 25% cell cost reduction, the deterministic cost per QALY gain ratio for ACI compared with MF was £10,523 for scenario 1 and £11,404 for scenario 2. The cost per QALY gain ratio for a 50% cell cost reduction for ACI compared with MF was £6651 (scenario 1) and £7210 (scenario 2), and the resulting figures for a 75% reduction were £2779 (scenario 1) and £3016 (scenario 2). With the reduction in cell costs, the cost-effectiveness of ACI improved relative to MF. Hence, the cost of cells was a key driver for the cost-effectiveness. These results were of similar magnitudes and directions for the probabilistic results.

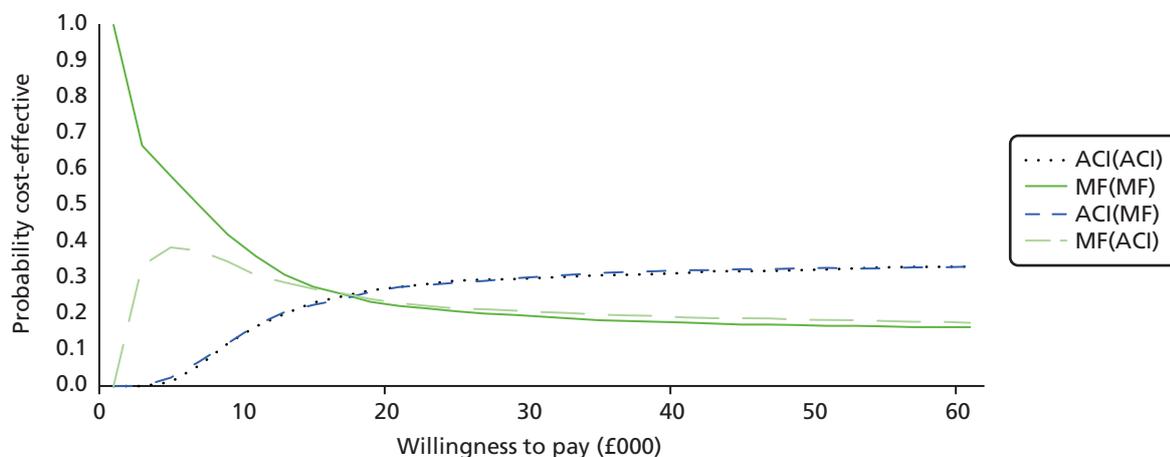


FIGURE 25 Cost-effectiveness acceptability curve: base-case results – all sequences.

TABLE 26 Sensitivity cost-effectiveness results (by scenario): cell cost reduction

| Procedure | Total mean costs (£) | Total mean QALYs | Incremental costs (£) | Incremental QALYs | ICER (£) (cost per QALY gained) |
|-------------------------------------|----------------------|------------------|-----------------------|-------------------|---------------------------------|
| Deterministic: 25% reduction | | | | | |
| <i>Scenario 1</i> | | | | | |
| MF(ACI) | 6183 | 17.0284 | – | – | – |
| ACI(ACI) | 16,647 | 18.0228 | 10,464 | 0.9944 | 10,523 |

continued

TABLE 26 Sensitivity cost-effectiveness results (by scenario): cell cost reduction (continued)

| Procedure | Total mean costs (£) | Total mean QALYs | Incremental costs (£) | Incremental QALYs | ICER (£) (cost per QALY gained) |
|-------------------------------------|----------------------|------------------|-----------------------|-------------------|---------------------------------|
| <i>Scenario 2</i> | | | | | |
| MF(MF) | 5015 | 17.0033 | – | – | – |
| ACI(MF) | 15,892 | 17.9570 | 10,877 | 0.9537 | 11,404 |
| Probabilistic: 25% reduction | | | | | |
| <i>Scenario 1</i> | | | | | |
| MF(ACI) | 6183 | 17.0305 | – | – | – |
| ACI(ACI) | 16,637 | 18.0497 | 10,454 | 1.0192 | 10,258 |
| <i>Scenario 2</i> | | | | | |
| MF(MF) | 5009 | 17.0086 | – | – | – |
| ACI(MF) | 15,880 | 17.9502 | 10,871 | 0.9416 | 11,545 |
| Deterministic: 50% reduction | | | | | |
| <i>Scenario 1</i> | | | | | |
| MF(ACI) | 5760 | 17.0284 | – | – | – |
| ACI(ACI) | 12,373 | 18.0228 | 6614 | 0.9944 | 6651 |
| <i>Scenario 2</i> | | | | | |
| MF(MF) | 5015 | 17.0033 | – | – | – |
| ACI(MF) | 11,892 | 17.9570 | 6877 | 0.9537 | 7210 |
| Probabilistic: 50% reduction | | | | | |
| <i>Scenario 1</i> | | | | | |
| MF(ACI) | 5770 | 17.0250 | – | – | – |
| ACI(ACI) | 12,362 | 18.0100 | 6592 | 0.9850 | 6693 |
| <i>Scenario 2</i> | | | | | |
| MF(MF) | 5020 | 16.9907 | – | – | – |
| ACI(MF) | 11,876 | 17.9123 | 6856 | 0.9216 | 7439 |
| Deterministic: 75% reduction | | | | | |
| <i>Scenario 1</i> | | | | | |
| MF(ACI) | 5336 | 17.0284 | – | – | – |
| ACI(ACI) | 8100 | 18.0228 | 2763 | 0.9944 | 2779 |
| <i>Scenario 2</i> | | | | | |
| MF(MF) | 5015 | 17.0033 | – | – | – |
| ACI(MF) | 7892 | 17.9570 | 2877 | 0.9537 | 3016 |
| Probabilistic: 75% reduction | | | | | |
| <i>Scenario 1</i> | | | | | |
| MF(ACI) | 5346 | 16.9755 | – | – | – |
| ACI(ACI) | 8083 | 18.0442 | 2737 | 1.0687 | 2561 |
| <i>Scenario 2</i> | | | | | |
| MF(MF) | 5023 | 16.9546 | – | – | – |
| ACI(MF) | 7878 | 17.9253 | 2854 | 0.9707 | 2940 |

Figure 26a–f presents the CEACs for the SA for cell cost reductions for scenarios 1 and 2. For a 25% cell cost reduction: scenario 1, if the decision-maker was willing to pay £20,000, the probability that ACI was more cost-effective than MF was 65% (see Figure 26a); scenario 2, the probability that ACI was more cost-effective than MF was 58% (see Figure 26b). For a 50% cell cost reduction: scenario 1, if the decision-maker was willing to pay £20,000, the probability that ACI was more cost-effective than MF was 67% (see Figure 26c); scenario 2, the probability that ACI was more cost-effective than MF was 64% (see Figure 26d). For a 75% cell cost reduction: for scenario 1, if the decision-maker was willing to pay £20,000, there was a 71% probability that ACI was more cost-effective than MF (see Figure 26e); scenario 2, there was a 70% probability that ACI was more cost-effective than MF (see Figure 26f). The graphs indicate that reductions in cell costs improve the cost-effectiveness of ACI compared with MF.

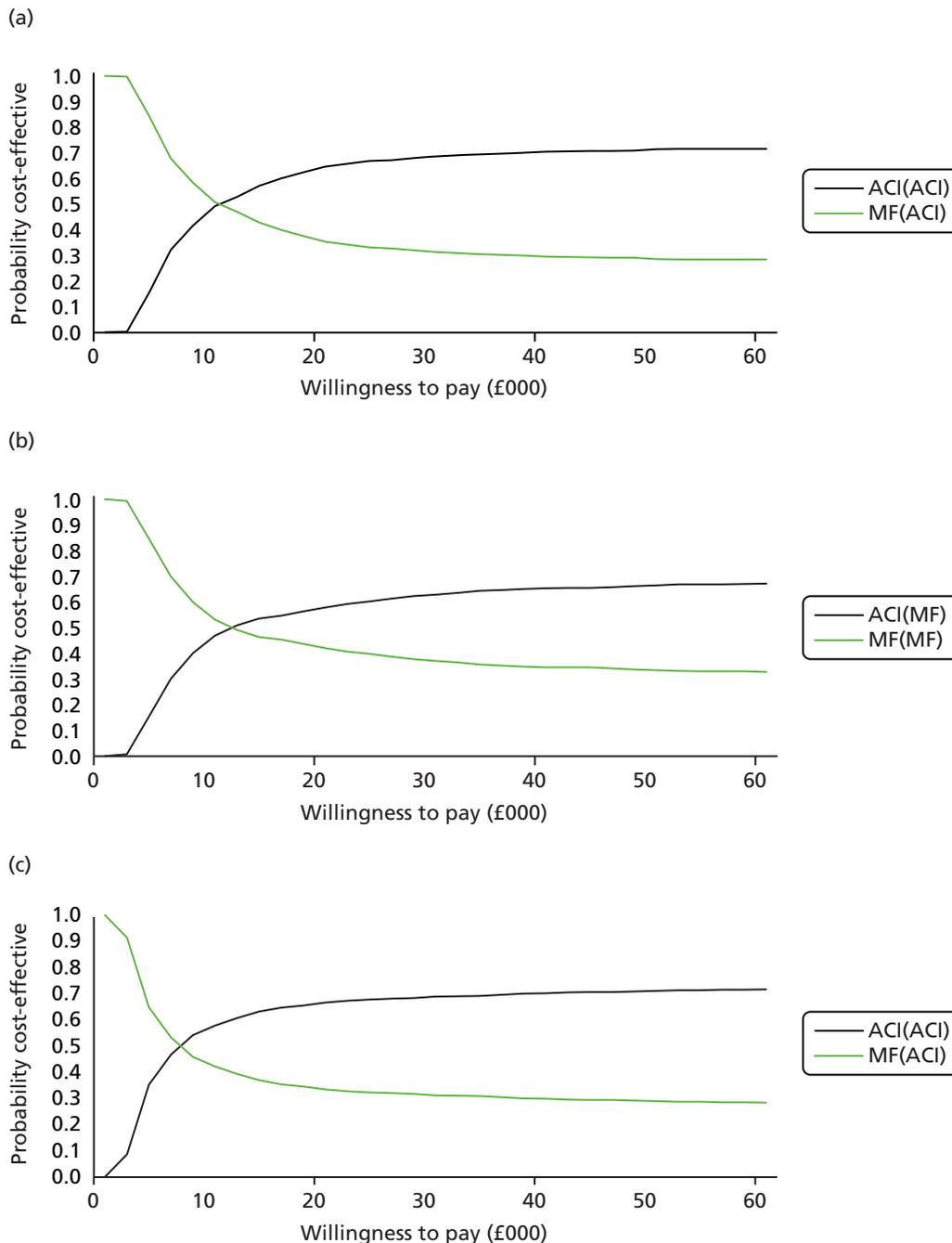


FIGURE 26 Cost-effectiveness acceptability curves: cost reductions: (a) 25% cell cost reduction – scenario 1; (b) 25% cell cost reduction – scenario 2; (c) 50% cell cost reduction – scenario 1; (d) 50% cell cost reduction – scenario 2; (e) 75% cell cost reduction – scenario 1; and (f) 75% cell cost reduction – scenario 2. (*continued*)

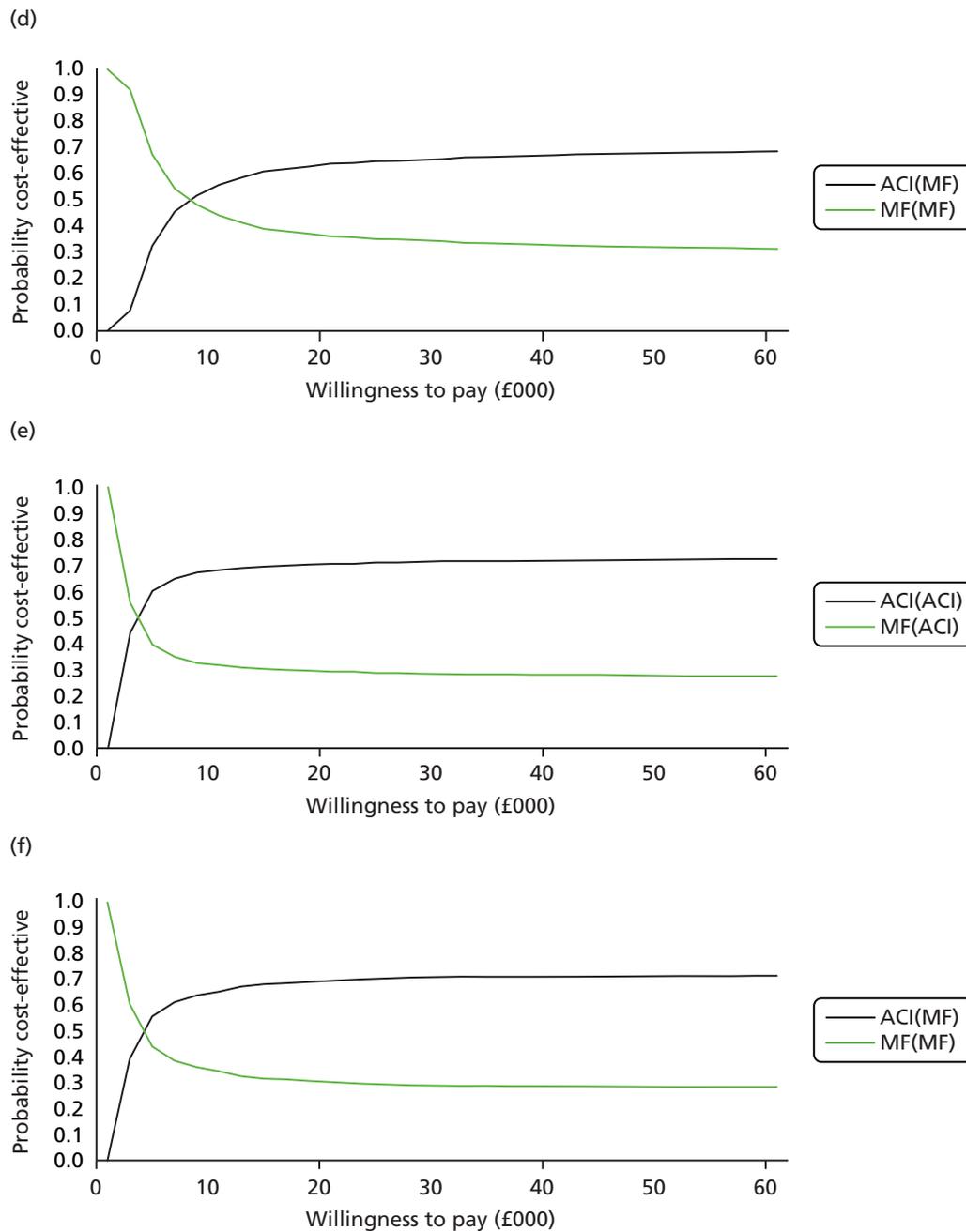


FIGURE 26 Cost-effectiveness acceptability curves: cost reductions: (a) 25% cell cost reduction – scenario 1; (b) 25% cell cost reduction – scenario 2; (c) 50% cell cost reduction – scenario 1; (d) 50% cell cost reduction – scenario 2; (e) 75% cell cost reduction – scenario 1; and (f) 75% cell cost reduction – scenario 2.

Table 27 shows the deterministic and probabilistic cost-effectiveness results for the lifetime horizon for cell cost reduction, and results were ranked by least costly option. When the cost of cells was reduced by 25%, 50% and 75%, these results were in line with the base-case cost-effectiveness results. That is, for the discounted deterministic results MF(MF) was the least costly option and had the fewest QALYs, whereas ACI(ACI) was the most expensive option, but generated more QALYs. Again, the option MF(ACI) was extendedly dominated by a linear combination of MF(MF) and ACI(MF) and therefore this option was eliminated from the comparison. The deterministic ICER between ACI(MF) and MF(MF) was just over £11,000; and the ICER between the two ACI options: ACI(ACI) vs. ACI(MF) was just under £11,500 when there was a 25% reduction in costs. These ICER figures are £7200 and £7300, respectively, when there was a 50% reduction in costs, and the corresponding figures are £3000 and £3200, respectively, when there was a 75% reduction in costs. For all cell cost reduction scenarios, both the deterministic and probabilistic

TABLE 27 Sensitivity cost-effectiveness results: cell cost reduction

| Procedure | Total mean costs (£) | Total mean QALYs | Comparison | Incremental costs (£) | Incremental QALYs | ICER (£) (cost per QALY gained) |
|-------------------------------------|----------------------|------------------|----------------------|-----------------------|---------------------|---------------------------------|
| Deterministic: 25% reduction | | | | | | |
| MF(MF) | 5015 | 17.0033 | – | – | – | – |
| MF(ACI) | 6183 | 17.0284 | MF(ACI) vs. MF(MF) | 1168 | 0.0251 ^a | Extendedly dominated |
| ACI(MF) | 15,892 | 17.9570 | ACI(MF) vs. MF(MF) | 10,877 | 0.9537 | 11,404 |
| ACI(ACI) | 16,647 | 18.0228 | ACI(ACI) vs. ACI(MF) | 755 | 0.0658 ^a | 11,483 |
| Probabilistic: 25% reduction | | | | | | |
| MF(MF) | 5009 | 17.0086 | – | – | – | – |
| MF(ACI) | 6183 | 17.0305 | MF(ACI) vs. MF(MF) | 1174 | 0.0219 ^a | Extendedly dominated |
| ACI(MF) | 15,880 | 17.9502 | ACI(MF) vs. MF(MF) | 10,871 | 0.9416 | 11,545 |
| ACI(ACI) | 16,637 | 18.0497 | ACI(ACI) vs. ACI(MF) | 758 | 0.0994 ^a | 7618 |
| Deterministic: 50% reduction | | | | | | |
| MF(MF) | 5015 | 17.0033 | – | – | – | – |
| MF(ACI) | 5760 | 17.0284 | MF(ACI) vs. MF(MF) | 744 | 0.0251 ^a | Extendedly dominated |
| ACI(MF) | 11,892 | 17.9570 | ACI(MF) vs. MF(MF) | 6877 | 0.9537 | 7210 |
| ACI(ACI) | 12,373 | 18.0228 | ACI(ACI) vs. ACI(MF) | 481 | 0.0658 ^a | 7319 |
| Probabilistic: 50% reduction | | | | | | |
| MF(MF) | 5020 | 16.9907 | – | – | – | – |
| MF(ACI) | 5770 | 17.0250 | MF(ACI) vs. MF(MF) | 750 | 0.0343 ^a | Extendedly dominated |
| ACI(MF) | 11,876 | 17.9123 | ACI(MF) vs. MF(MF) | 6856 | 0.9216 | 7439 |
| ACI(ACI) | 12,362 | 18.0100 | ACI(ACI) vs. ACI(MF) | 486 | 0.0977 ^a | 4979 |
| Deterministic: 75% reduction | | | | | | |
| MF(MF) | 5015 | 17.0033 | – | – | – | – |
| MF(ACI) | 5336 | 17.0284 | MF(ACI) vs. MF(MF) | 321 | 0.0251 ^a | Extendedly dominated |
| ACI(MF) | 7892 | 17.9570 | ACI(MF) vs. MF(MF) | 2877 | 0.9537 | 3016 |
| ACI(ACI) | 8100 | 18.0228 | ACI(ACI) vs. ACI(MF) | 207 | 0.0658 ^a | 3155 |
| Probabilistic: 75% reduction | | | | | | |
| MF(MF) | 5023 | 16.9546 | – | – | – | – |
| MF(ACI) | 5346 | 16.9755 | MF(ACI) vs. MF(MF) | 322 | 0.0209 ^a | Extendedly dominated |
| ACI(MF) | 7878 | 17.9253 | ACI(MF) vs. MF(MF) | 2854 | 0.9707 | 2940 |
| ACI(ACI) | 8083 | 18.0442 | ACI(ACI) vs. ACI(MF) | 205 | 0.1189 ^a | 1725 |

^a As the incremental QALYs are near zero, the ICER can fluctuate widely.

results indicate that ACI as a first procedure was more cost-effective than MF as a first procedure, and from these results, again, we see that the cost of cells is a key driver of the cost-effectiveness estimates.

Figure 27a–c presents the CEACs for the SA results for the cell cost reduction. The graphs clearly show that if the decision-maker is willing to pay £20,000 then the probability that ACI(ACI) is more cost-effective than the other three comparisons is 32% for a 50% reduction in the costs of cells (although there is not

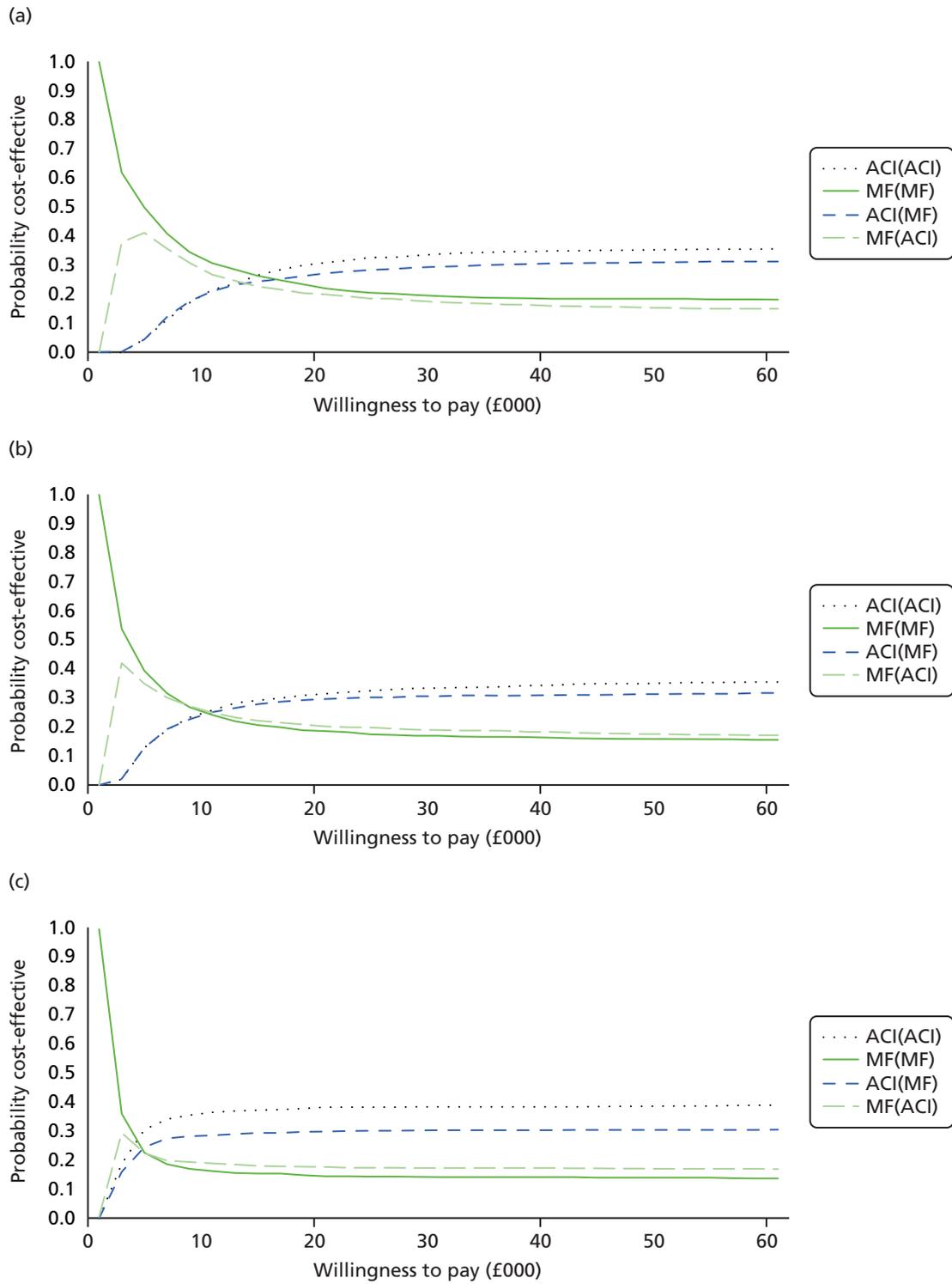


FIGURE 27 Cost-effectiveness acceptability curves: cell cost reduction: (a) 25% cell cost reduction; (b) 50% cell cost reduction; and (c) 75% cell cost reduction.

much difference if MF was the second repair after the ACI) and 38% for a 75% reduction in the costs of cells. Whereas, if the decision-maker pays £30,000 for a 25% reduction in the cost of cells then the probability that ACI(ACI) is more cost-effective than the other three comparisons is 34%. This suggests that ACI as first procedure is more cost-effective than MF as first procedure.

Sensitivity analysis 2: changing the time horizon

In the base-case analysis, a lifetime horizon was chosen with the starting age 33 years for the cohort. In this SA we have varied the time horizon (10, 20, 30, 40 and 50 years) to see how this affects the ICER. Table 28 shows the sensitivity cost-effectiveness results for the different time horizons. For all of the time horizons, even though ACI was more costly than MF, there were more QALYs gained with ACI than MF. For the 10-year time horizon, the deterministic cost per QALY gain for ACI compared with MF was

TABLE 28 Sensitivity cost-effectiveness results (by scenario): changing the time horizon

| Procedure | Total mean costs (£) | Total mean QALYs | Incremental costs (£) | Incremental QALYs | ICER (£) (cost per QALY gained) |
|--|----------------------|------------------|-----------------------|-------------------|---------------------------------|
| Deterministic: 10-year time horizon | | | | | |
| <i>Scenario 1</i> | | | | | |
| MF(ACI) | 5983 | 7.3030 | – | – | – |
| ACI(ACI) | 20,082 | 7.8454 | 14,098 | 0.5424 | 25,992 |
| <i>Scenario 2</i> | | | | | |
| MF(MF) | 4498 | 7.2906 | – | – | – |
| ACI(MF) | 19,329 | 7.8321 | 14,831 | 0.5415 | 27,388 |
| Probabilistic: 10-year time horizon | | | | | |
| <i>Scenario 1</i> | | | | | |
| MF(ACI) | 5989 | 7.2950 | – | – | – |
| ACI(ACI) | 20,075 | 7.8501 | 14,086 | 0.5550 | 25,379 |
| <i>Scenario 2</i> | | | | | |
| MF(MF) | 4505 | 7.2845 | – | – | – |
| ACI(MF) | 19,326 | 7.8270 | 14,821 | 0.5425 | 27,320 |
| Deterministic: 20-year time horizon | | | | | |
| <i>Scenario 1</i> | | | | | |
| MF(ACI) | 6104 | 11.2812 | – | – | – |
| ACI(ACI) | 20,340 | 12.1040 | 14,286 | 0.8228 | 17,301 |
| <i>Scenario 2</i> | | | | | |
| MF(MF) | 4524 | 11.2587 | – | – | – |
| ACI(MF) | 19,384 | 12.0654 | 14,860 | 0.8067 | 18,421 |
| Probabilistic: 20-year time horizon | | | | | |
| <i>Scenario 1</i> | | | | | |
| MF(ACI) | 6098 | 11.2630 | – | – | – |
| ACI(ACI) | 20,359 | 12.1136 | 14,261 | 0.8506 | 16,766 |
| <i>Scenario 2</i> | | | | | |
| MF(MF) | 4526 | 11.2362 | – | – | – |
| ACI(MF) | 19,414 | 12.0625 | 14,888 | 0.8263 | 18,019 |
| Deterministic: 30-year time horizon | | | | | |
| <i>Scenario 1</i> | | | | | |
| MF(ACI) | 6329 | 13.9997 | – | – | – |
| ACI(ACI) | 20,614 | 14.9318 | 14,285 | 0.9321 | 15,326 |

continued

TABLE 28 Sensitivity cost-effectiveness results (by scenario): changing the time horizon (*continued*)

| Procedure | Total mean costs (£) | Total mean QALYs | Incremental costs (£) | Incremental QALYs | ICER (£) (cost per QALY gained) |
|--|----------------------|------------------|-----------------------|-------------------|---------------------------------|
| <i>Scenario 2</i> | | | | | |
| MF(MF) | 4739 | 13.9750 | – | – | – |
| ACI(MF) | 19,609 | 14.8774 | 14,871 | 0.9024 | 16,480 |
| Probabilistic: 30-year time horizon | | | | | |
| <i>Scenario 1</i> | | | | | |
| MF(ACI) | 6326 | 14.0117 | – | – | – |
| ACI(ACI) | 20,642 | 14.9494 | 14,316 | 0.9377 | 15,267 |
| <i>Scenario 2</i> | | | | | |
| MF(MF) | 4728 | 13.9748 | – | – | – |
| ACI(MF) | 19,628 | 14.8754 | 14,900 | 0.9006 | 16,545 |
| Deterministic: 40-year time horizon | | | | | |
| <i>Scenario 1</i> | | | | | |
| MF(ACI) | 6492 | 15.7604 | – | – | – |
| ACI(ACI) | 20,798 | 16.7368 | 14,306 | 0.9764 | 14,652 |
| <i>Scenario 2</i> | | | | | |
| MF(MF) | 4901 | 15.7354 | – | – | – |
| ACI(MF) | 19,775 | 16.6747 | 14,875 | 0.9393 | 15,836 |
| Probabilistic: 40-year time horizon | | | | | |
| <i>Scenario 1</i> | | | | | |
| MF(ACI) | 6494 | 15.7558 | – | – | – |
| ACI(ACI) | 20,763 | 16.7219 | 14,269 | 0.9662 | 14,768 |
| <i>Scenario 2</i> | | | | | |
| MF(MF) | 4900 | 15.7279 | – | – | – |
| ACI(MF) | 19,376 | 16.6504 | 14,837 | 0.9225 | 16,083 |
| Deterministic: 50-year time horizon | | | | | |
| <i>Scenario 1</i> | | | | | |
| MF(ACI) | 6579 | 16.7164 | – | – | – |
| ACI(ACI) | 20,891 | 17.7078 | 14,313 | 0.9914 | 14,437 |
| <i>Scenario 2</i> | | | | | |
| MF(MF) | 4987 | 16.6913 | – | – | – |
| ACI(MF) | 19,864 | 17.6427 | 14,876 | 0.9514 | 15,636 |
| Probabilistic: 50-year time horizon | | | | | |
| <i>Scenario 1</i> | | | | | |
| MF(ACI) | 6557 | 16.7119 | – | – | – |
| ACI(ACI) | 20,841 | 17.6964 | 14,284 | 0.9845 | 14,509 |
| <i>Scenario 2</i> | | | | | |
| MF(MF) | 4974 | 16.6777 | – | – | – |
| ACI(MF) | 19,820 | 17.6182 | 14,845 | 0.9405 | 15,785 |

£25,992 for scenario 1 and £27,388 for scenario 2. The cost per QALY gained for the two scenarios ranged from £17,000 to £18,000 for the 20-year time horizon; £15,000 to £16,000 for the 30- and 40-year time horizons; and £14–16,000 for the 50-year time horizon. For both scenarios, ACI as a first repair was more-cost-effective than MF as a first repair, the longer the time horizon. These results were of similar magnitudes and directions for the probabilistic results.

Figure 28a–j presents the CEACs for the SA for the different time horizons for scenarios 1 and 2. For the 10-year time horizon, scenario 1: if the decision-maker was willing to pay £20,000 then the probability that MF was more cost-effective than ACI was 62% (see Figure 28a); for scenario 2, MF was 64% more

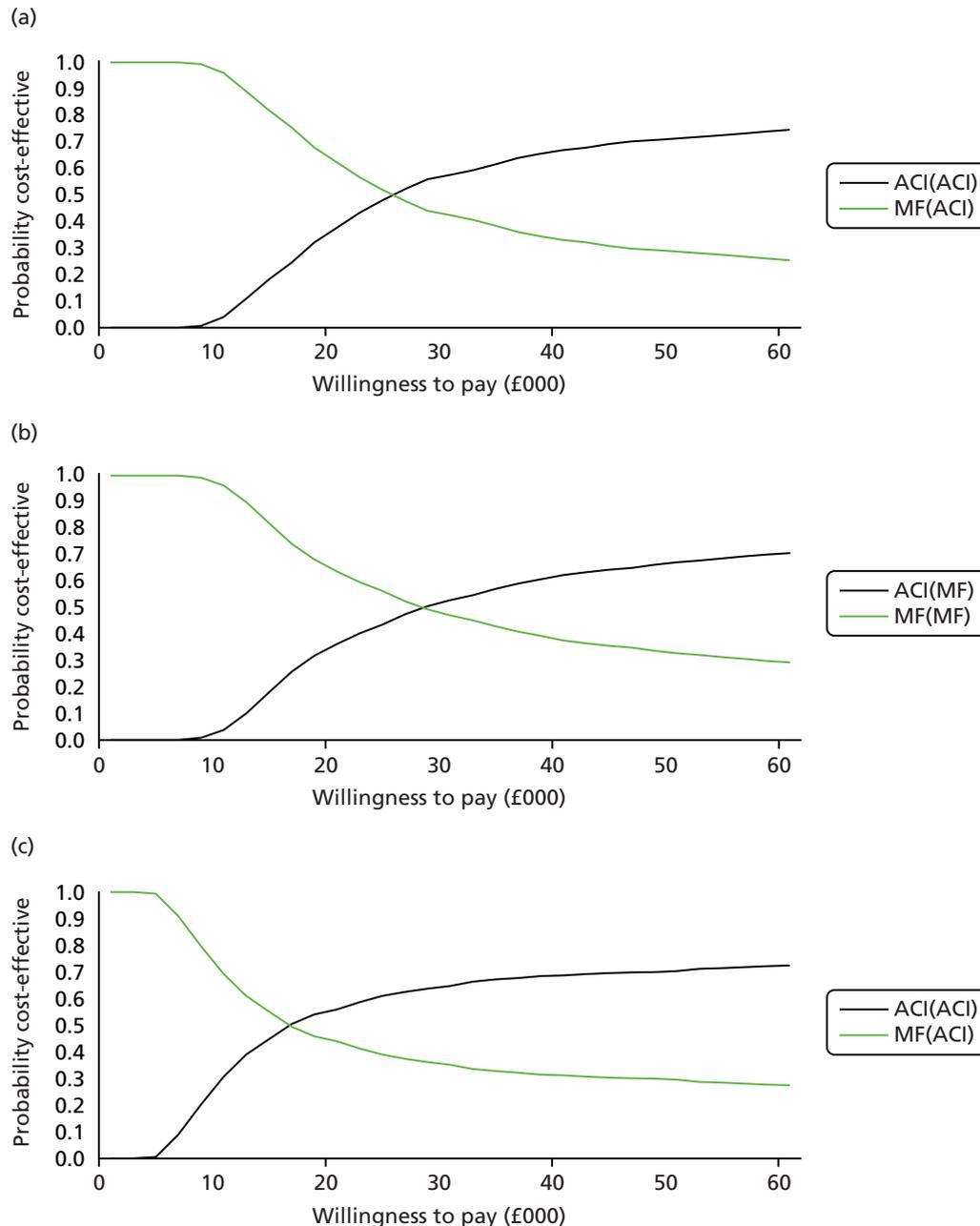


FIGURE 28 Cost-effectiveness acceptability curves: different time horizons: (a) 10-year time horizon – scenario 1; (b) 10-year time horizon – scenario 2; (c) 20-year time horizon – scenario 1; (d) 20-year time horizon – scenario 2; (e) 30-year time horizon – scenario 1; (f) 30-year time horizon – scenario 2; (g) 40-year time horizon – scenario 1; (h) 40-year time horizon – scenario 2; (i) 50-year time horizon – scenario 1; and (j) 50-year time horizon – scenario 2. (continued)

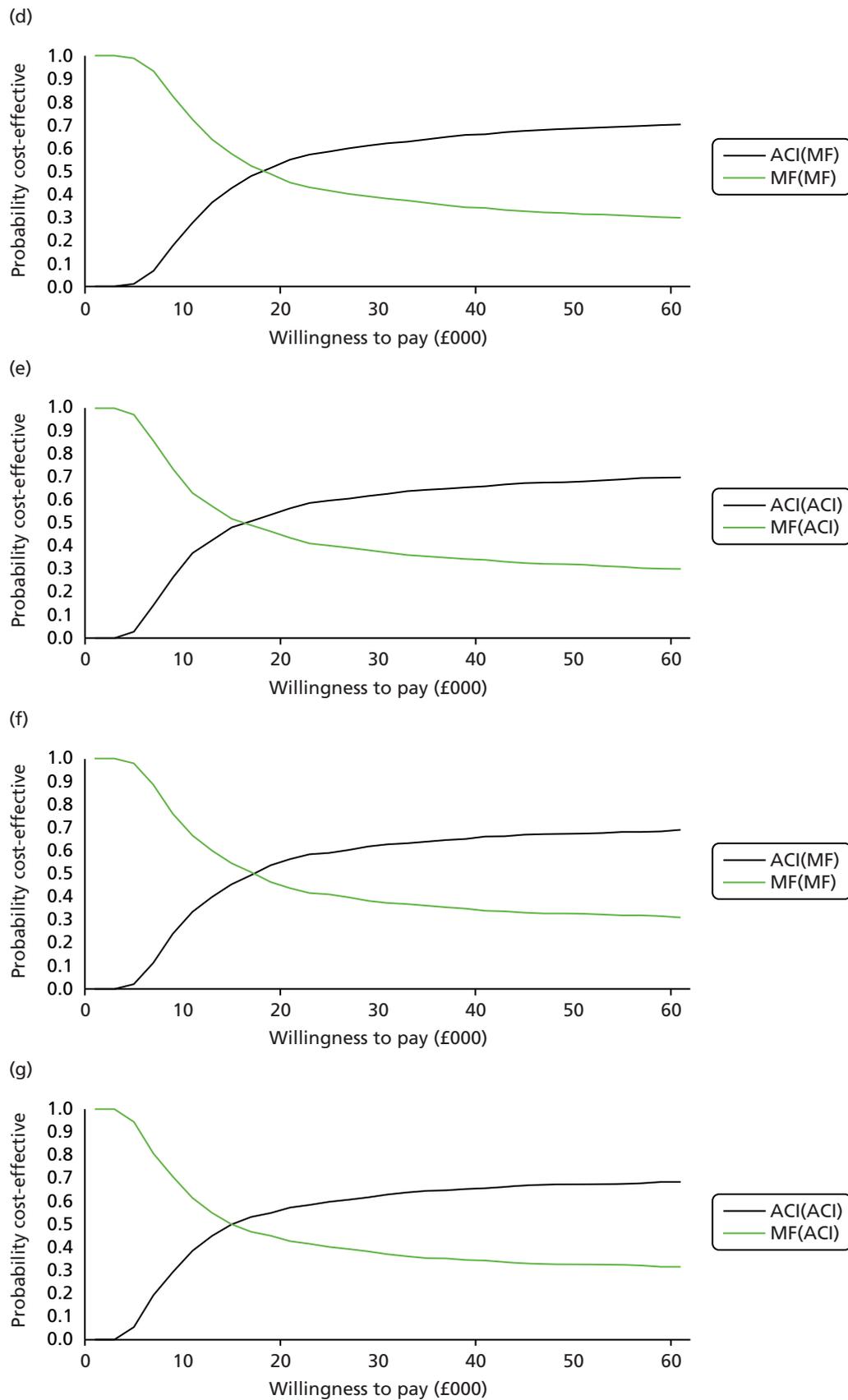


FIGURE 28 Cost-effectiveness acceptability curves: different time horizons: (a) 10-year time horizon – scenario 1; (b) 10-year time horizon – scenario 2; (c) 20-year time horizon – scenario 1; (d) 20-year time horizon – scenario 2; (e) 30-year time horizon – scenario 1; (f) 30-year time horizon – scenario 2; (g) 40-year time horizon – scenario 1; (h) 40-year time horizon – scenario 2; (i) 50-year time horizon – scenario 1; and (j) 50-year time horizon – scenario 2. (continued)

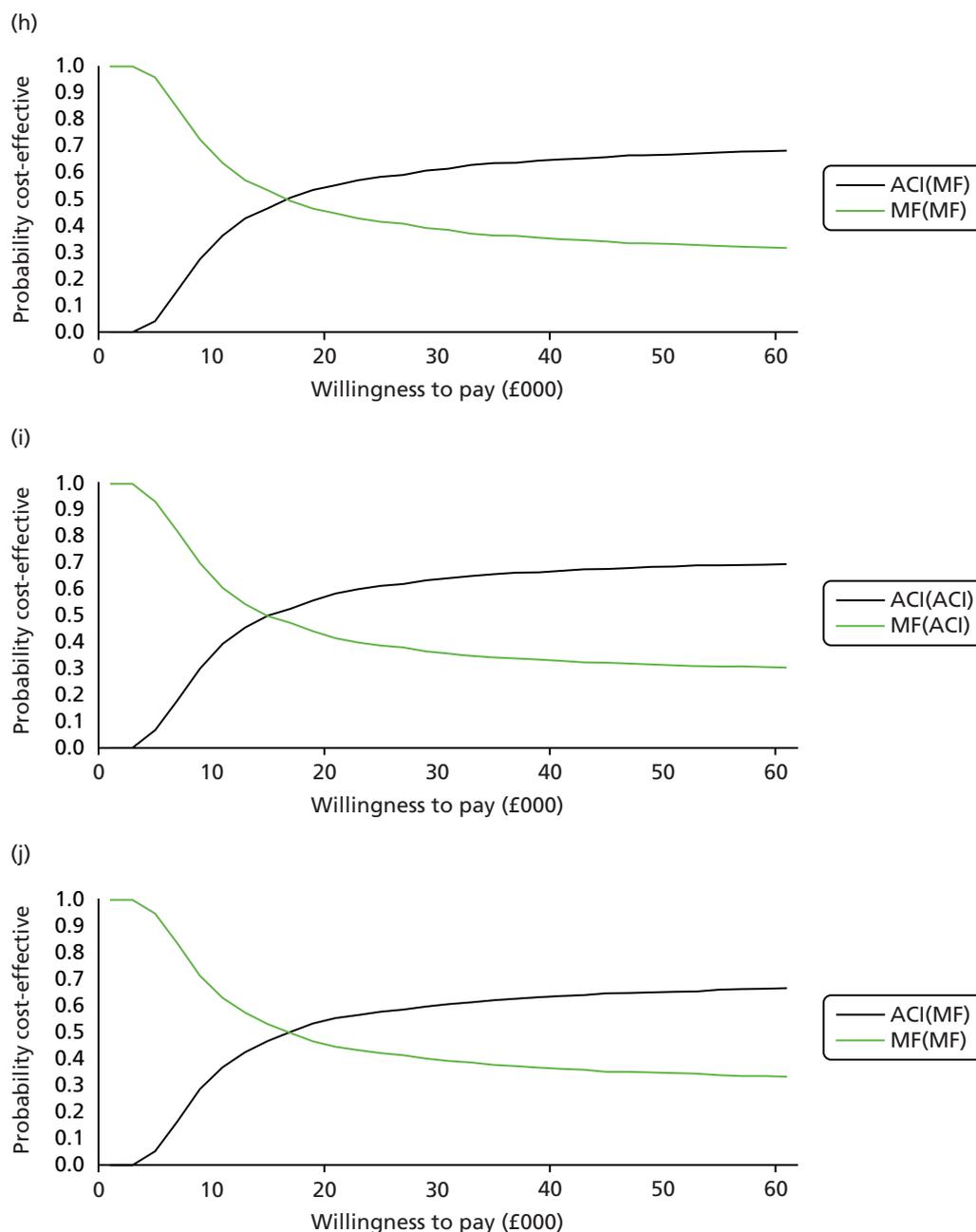


FIGURE 28 Cost-effectiveness acceptability curves: different time horizons: (a) 10-year time horizon – scenario 1; (b) 10-year time horizon – scenario 2; (c) 20-year time horizon – scenario 1; (d) 20-year time horizon – scenario 2; (e) 30-year time horizon – scenario 1; (f) 30-year time horizon – scenario 2; (g) 40-year time horizon – scenario 1; (h) 40-year time horizon – scenario 2; (i) 50-year time horizon – scenario 1; and (j) 50-year time horizon – scenario 2.

likely to be cost-effective than ACI (see *Figure 28b*). ACI became more cost-effective than MF when the decision-maker was willing to pay approximately £26,000 for scenario 1 and £28,000 for scenario 2. For all other time horizons, the probability that ACI was more cost-effective than MF was approximately 55% for both scenarios when the decision-maker was willing to pay £20,000. The results highlighted that, for the longer time horizons, ACI as a first repair was more cost-effective than MF as a first repair.

Table 29 shows the deterministic and probabilistic cost-effectiveness results for the different time horizons and results were ranked by the least costly option. Again, the option MF(ACI) was extendedly dominated by a linear combination of MF(MF) and ACI(MF) and therefore this option was eliminated from the

TABLE 29 Sensitivity cost-effectiveness results: changing the time horizon

| Procedure | Total mean costs (£) | Total mean QALYs | Comparison | Incremental costs (£) | Incremental QALYs | ICER (£) (cost per QALY gained) |
|--|----------------------|------------------|----------------------|-----------------------|---------------------|---------------------------------|
| Deterministic: 10-year time horizon | | | | | | |
| MF(MF) | 4498 | 7.2906 | – | – | – | – |
| MF(ACI) | 5983 | 7.3030 | MF(ACI) vs. MF(MF) | 1485 | 0.0124 ^a | Extendedly dominated |
| ACI(MF) | 19,329 | 7.8321 | ACI(MF) vs. MF(MF) | 14,831 | 0.5415 | 27,388 |
| ACI(ACI) | 20,082 | 7.8454 | ACI(ACI) vs. ACI(MF) | 753 | 0.0132 ^a | 56,816 |
| Probabilistic: 10-year time horizon | | | | | | |
| MF(MF) | 4505 | 7.2845 | – | – | – | – |
| MF(ACI) | 5989 | 7.2950 | MF(ACI) vs. MF(MF) | 1484 | 0.0105 ^a | Extendedly dominated |
| ACI(MF) | 19,326 | 7.8270 | ACI(MF) vs. MF(MF) | 14,821 | 0.5425 | 27,320 |
| ACI(ACI) | 20,075 | 7.8501 | ACI(ACI) vs. ACI(MF) | 749 | 0.0231 ^a | 32,448 |
| Deterministic: 20-year time horizon | | | | | | |
| MF(MF) | 4524 | 11.2587 | – | – | – | – |
| MF(ACI) | 6104 | 11.2812 | MF(ACI) vs. MF(MF) | 1580 | 0.0225 ^a | Extendedly dominated |
| ACI(MF) | 19,384 | 12.0654 | ACI(MF) vs. MF(MF) | 14,860 | 0.8067 | 18,421 |
| ACI(ACI) | 20,340 | 12.1040 | ACI(ACI) vs. ACI(MF) | 955 | 0.0386 ^a | 24,742 |
| Probabilistic: 20-year time horizon | | | | | | |
| MF(MF) | 4526 | 11.2362 | – | – | – | – |
| MF(ACI) | 6098 | 11.2630 | MF(ACI) vs. MF(MF) | 1572 | 0.0268 ^a | Extendedly dominated |
| ACI(MF) | 19,414 | 12.0625 | ACI(MF) vs. MF(MF) | 14,888 | 0.8263 | 18,019 |
| ACI(ACI) | 20,359 | 12.1136 | ACI(ACI) vs. ACI(MF) | 944 | 0.0511 ^a | 18,486 |
| Deterministic: 30-year time horizon | | | | | | |
| MF(MF) | 4739 | 13.9750 | – | – | – | – |
| MF(ACI) | 6329 | 13.9997 | MF(ACI) vs. MF(MF) | 1590 | 0.0247 ^a | Extendedly dominated |
| ACI(MF) | 19,609 | 14.8774 | ACI(MF) vs. MF(MF) | 14,871 | 0.9024 | 16,480 |
| ACI(ACI) | 20,614 | 14.9318 | ACI(ACI) vs. ACI(MF) | 1005 | 0.0544 ^a | 18,472 |
| Probabilistic: 30-year time horizon | | | | | | |
| MF(MF) | 4728 | 13.9748 | – | – | – | – |
| MF(ACI) | 6326 | 14.0117 | MF(ACI) vs. MF(MF) | 1598 | 0.0369 ^a | Extendedly dominated |
| ACI(MF) | 19,628 | 14.8754 | ACI(MF) vs. MF(MF) | 14,900 | 0.9006 | 16,545 |
| ACI(ACI) | 20,642 | 14.9494 | ACI(ACI) vs. ACI(MF) | 1014 | 0.0740 ^a | 13,705 |
| Deterministic: 40-year time horizon | | | | | | |
| MF(MF) | 4901 | 15.7354 | – | – | – | – |
| MF(ACI) | 6492 | 15.7604 | MF(ACI) vs. MF(MF) | 1591 | 0.0250 ^a | Extendedly dominated |
| ACI(MF) | 19,775 | 16.6747 | ACI(MF) vs. MF(MF) | 14,875 | 0.9393 | 15,836 |
| ACI(ACI) | 20,798 | 16.7368 | ACI(ACI) vs. ACI(MF) | 1022 | 0.0621 ^a | 16,466 |

TABLE 29 Sensitivity cost-effectiveness results: changing the time horizon (*continued*)

| Procedure | Total mean costs (£) | Total mean QALYs | Comparison | Incremental costs (£) | Incremental QALYs | ICER (£) (cost per QALY gained) |
|--|----------------------|------------------|----------------------|-----------------------|---------------------|---------------------------------|
| Probabilistic: 40-year time horizon | | | | | | |
| MF(MF) | 4900 | 15.7279 | – | – | – | – |
| MF(ACI) | 6494 | 15.7558 | MF(ACI) vs. MF(MF) | 1594 | 0.0276 ^a | Extendedly dominated |
| ACI(MF) | 19,736 | 16.6504 | ACI(MF) vs. MF(MF) | 14,837 | 0.9225 | 16,083 |
| ACI(ACI) | 20,763 | 16.7219 | ACI(ACI) vs. ACI(MF) | 1026 | 0.0716 ^a | 14,336 |
| Deterministic: 50-year time horizon | | | | | | |
| MF(MF) | 4987 | 16.6913 | – | – | – | – |
| MF(ACI) | 6579 | 16.7164 | MF(ACI) vs. MF(MF) | 1592 | 0.0251 ^a | Extendedly dominated |
| ACI(MF) | 19,864 | 17.6427 | ACI(MF) vs. MF(MF) | 14,876 | 0.9514 | 15,636 |
| ACI(ACI) | 20,891 | 17.7078 | ACI(ACI) vs. ACI(MF) | 1028 | 0.0651 ^a | 15,793 |
| Probabilistic: 50-year time horizon | | | | | | |
| MF(MF) | 4974 | 16.6777 | – | – | – | – |
| MF(ACI) | 6557 | 16.7119 | MF(ACI) vs. MF(MF) | 1583 | 0.0341 ^a | Extendedly dominated |
| ACI(MF) | 19,820 | 17.6182 | ACI(MF) vs. MF(MF) | 14,845 | 0.9405 | 15,785 |
| ACI(ACI) | 20,841 | 17.6964 | ACI(ACI) vs. ACI(MF) | 1022 | 0.0782 ^a | 13,073 |

^a As the incremental QALYs are near zero, the ICER can fluctuate widely.

comparison. When comparing ACI(MF) with MF(MF) the deterministic ICER for a 10-year time horizon was over £27,000. For the same time period the deterministic ICER for the two initial ACI options: ACI(ACI) vs. ACI(MF) was approximately double this at £57,000. For the two initial ACI options, the deterministic ICER falls to just under £25,000 for a 20-year time horizon, and for the 30-year time horizon the ICER is just under £19,000. For both the 40- and 50-year time horizons (for the two initial ACI options), the deterministic ICER is very similar to the base-case ICER. The clear reason why the shorter time horizons are not cost-effective is due to the costs of ACI occurring at the start of the model and the benefits appearing much later, especially in terms of the reduced need for TKRs and fewer people going to the 'no further repair' or 'no further knee replacement' health states (where the utility is lower).

Figure 29a–e presents the CEACs for the SA results for the different time horizons. The graphs suggest that for the 10-year time horizon, MF(MF) is the most cost-effective option if the decision-maker is willing to pay £30,000 per QALY. At over £36,000 per QALY, the most cost-effective option is ACI as a first repair – if a second repair is needed then this could either be ACI or MF. For the 20-year time horizon, MF(MF) appears to be the most cost-effective option if the decision-maker is willing to pay £22,000 per QALY. At over £26,000 per QALY, the most cost-effective option is ACI(ACI) – ACI as a first repair, and, if a second repair is needed, this should also be ACI. As the time horizon increases ACI(ACI) – ACI as a first repair, and, if a second repair is needed, this should also be an ACI – is probably more cost-effective than the other three sequences. So, for example, if the decision-maker is willing to pay £24,000 per QALY then the probability that ACI(ACI) is more cost-effective for both the 30- and 40-year time horizons is 28%, and the probability that ACI(ACI) is more cost-effective for the 50-year time horizon is 29%.

Sensitivity analysis 3: microfracture done as a day case procedure

In the base-case analysis, according to clinical advice we have used a cost for MF as an inpatient procedure (£3020); however, we know that sometimes this procedure is done as day case. In the SA we have assumed that MF is done as a day case procedure at a cost of £1034. Table 30 shows the sensitivity

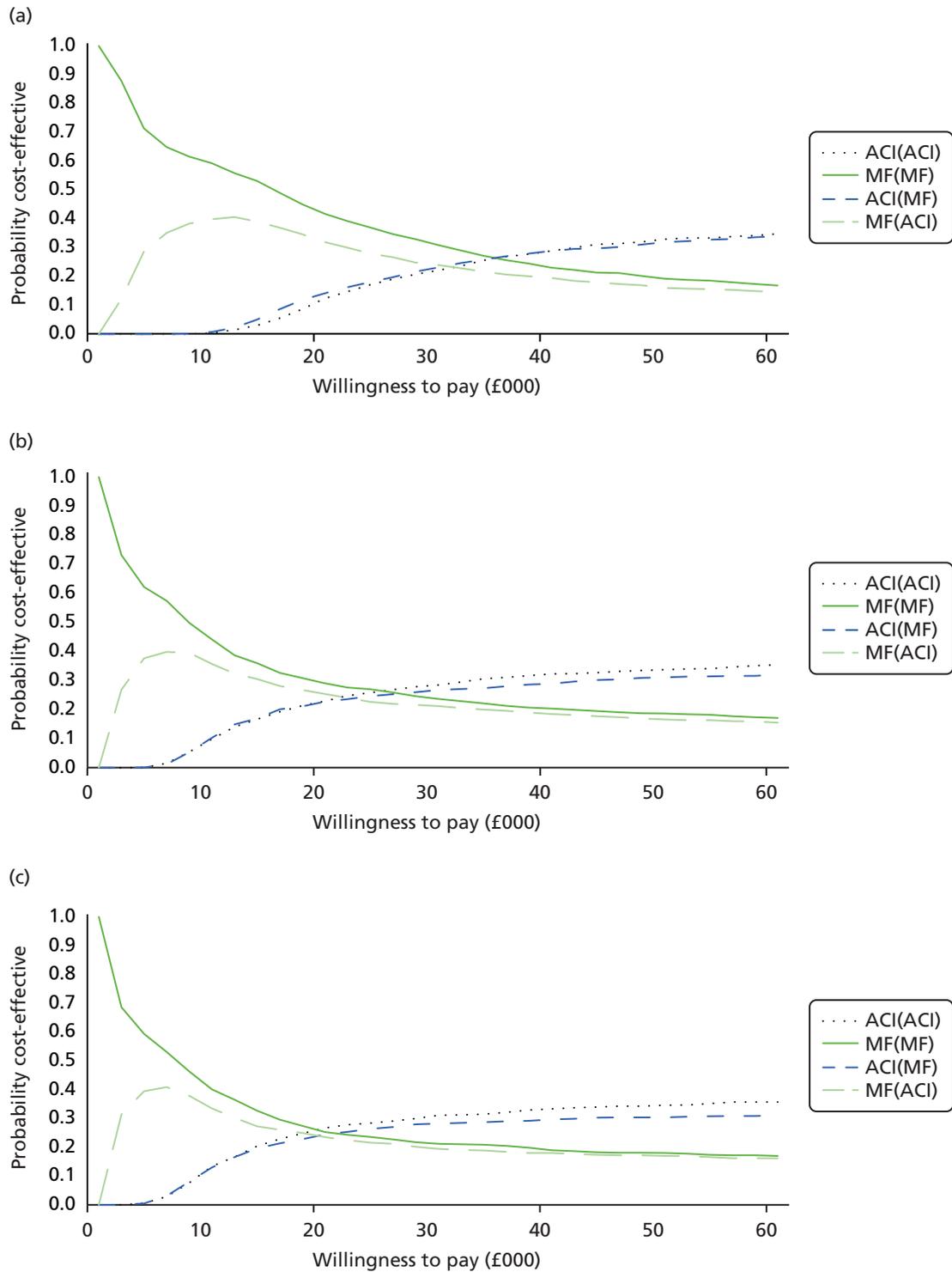


FIGURE 29 Cost-effectiveness acceptability curve: changing the time horizon. (a) 10-year time horizon; (b) 20-year time horizon; (c) 30-year time horizon; (d) 40-year time horizon; and (e) 50-year time horizon. (continued)

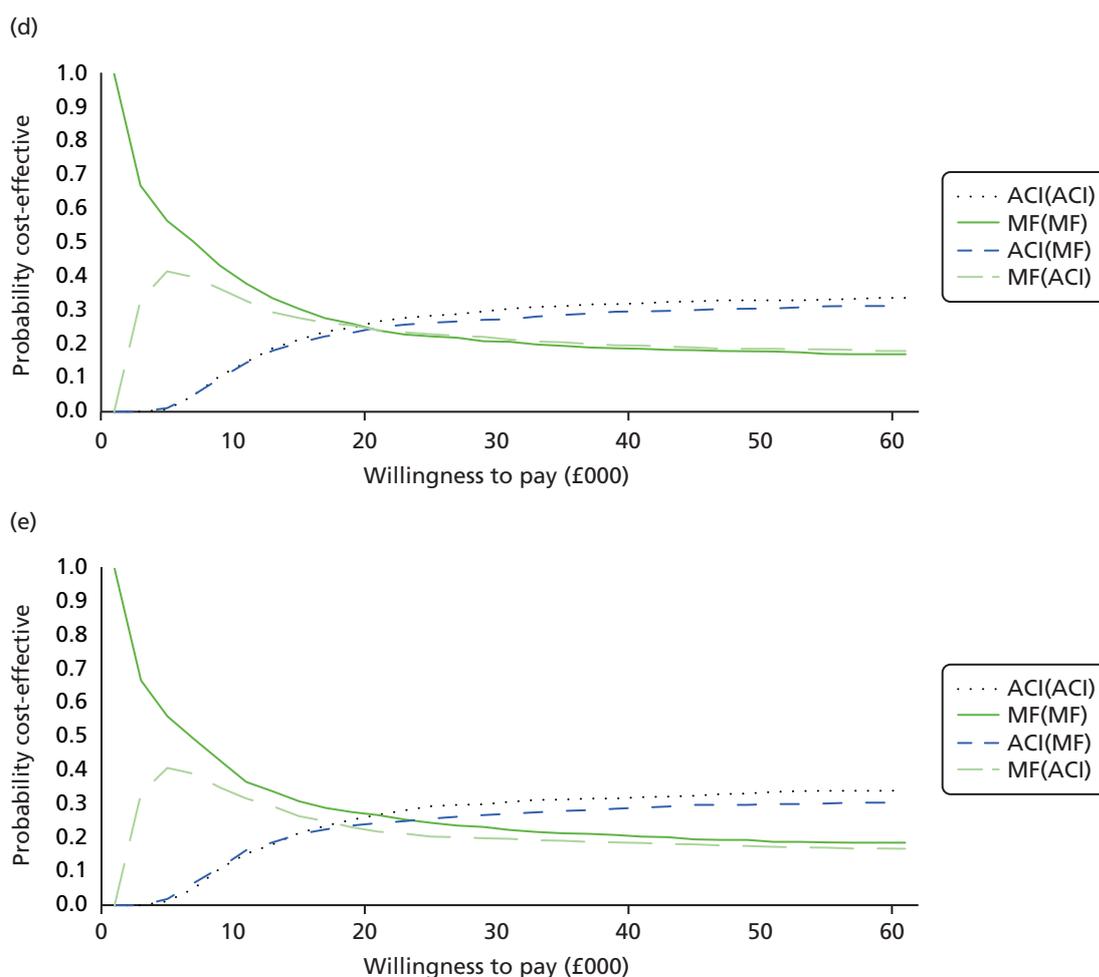


FIGURE 29 Cost-effectiveness acceptability curve: changing the time horizon. (a) 10-year time horizon; (b) 20-year time horizon; (c) 30-year time horizon; (d) 40-year time horizon; and (e) 50-year time horizon.

TABLE 30 Sensitivity cost-effectiveness results (by scenario): MF procedure as a day case surgery and not as an inpatient

| Procedure | Total mean costs (£) | Total mean QALYs | Incremental costs (£) | Incremental QALYs | ICER (£) (cost per QALY gained) |
|----------------------|----------------------|------------------|-----------------------|-------------------|---------------------------------|
| Deterministic | | | | | |
| <i>Scenario 1</i> | | | | | |
| MF(ACI) | 4621 | 17.0284 | – | – | – |
| ACI(ACI) | 20,921 | 18.0228 | 16,300 | 0.9944 | 16,391 |
| <i>Scenario 2</i> | | | | | |
| MF(MF) | 2819 | 17.0033 | – | – | – |
| ACI(MF) | 19,756 | 17.9570 | 16,937 | 0.9537 | 17,758 |
| Probabilistic | | | | | |
| <i>Scenario 1</i> | | | | | |
| MF(ACI) | 4620 | 17.0412 | – | – | – |
| ACI(ACI) | 20,951 | 17.9975 | 16,332 | 0.9563 | 17,078 |
| <i>Scenario 2</i> | | | | | |
| MF(MF) | 2811 | 17.0137 | – | – | – |
| ACI(MF) | 19,788 | 17.9065 | 16,977 | 0.8928 | 19,017 |

cost-effectiveness results for MF as a day case procedure. The costs for MF have fallen, but the QALY gain does not change. Hence, ACI as a first repair is still the most cost-effective procedure compared with MF as a first repair, with ICERs of just over £16,000 (scenario 1) and just under £18,000 (scenario 2).

Figure 30a and b presents the CEACs for MF as a day case procedure for scenarios 1 and 2, respectively. If the decision-maker was willing to pay £20,000, for scenario 1 the probability that ACI was more cost-effective than MF was 55% (see Figure 30a), and for scenario 2 the probability that ACI was more cost-effective than MF was 52% (see Figure 30b).

Table 31 presents the deterministic and probabilistic cost-effectiveness results. Compared with the base-case analysis, even though the costs for MF have fallen, there is no change in the QALYs.

The ICERs between the different options were in line with the base-case results and initial ACI appears more cost-effective than initial MF, and, for those who need a second repair after the first ACI, this should be another ACI.

Figure 31 presents the CEAC and the graph highlights that if the decision-maker is not willing to pay more than £22,000 then MF(MF) is the most cost-effective option. If willing to pay more than £24,000, then ACI(MF) is the most cost-effective option – that is, the first repair should be ACI and if a second repair is needed this should be MF because of the lower costs (even though having an ACI as a second repair generates more QALYs).

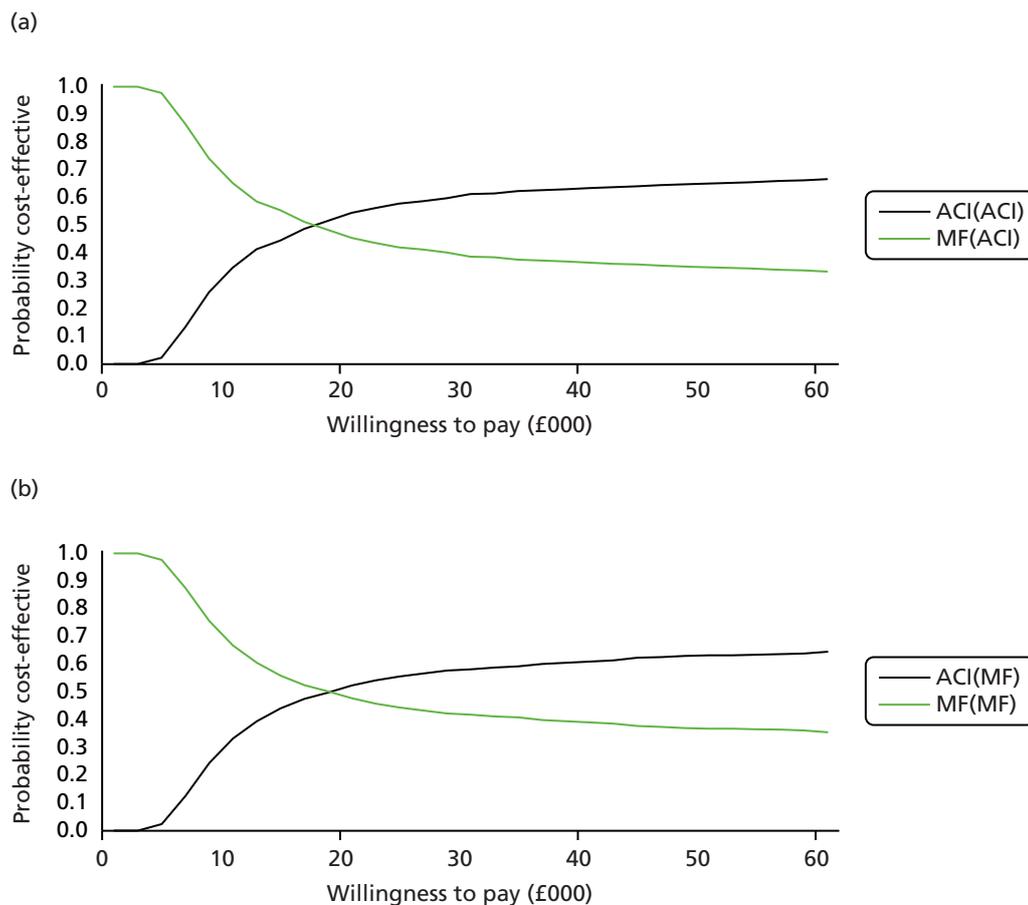
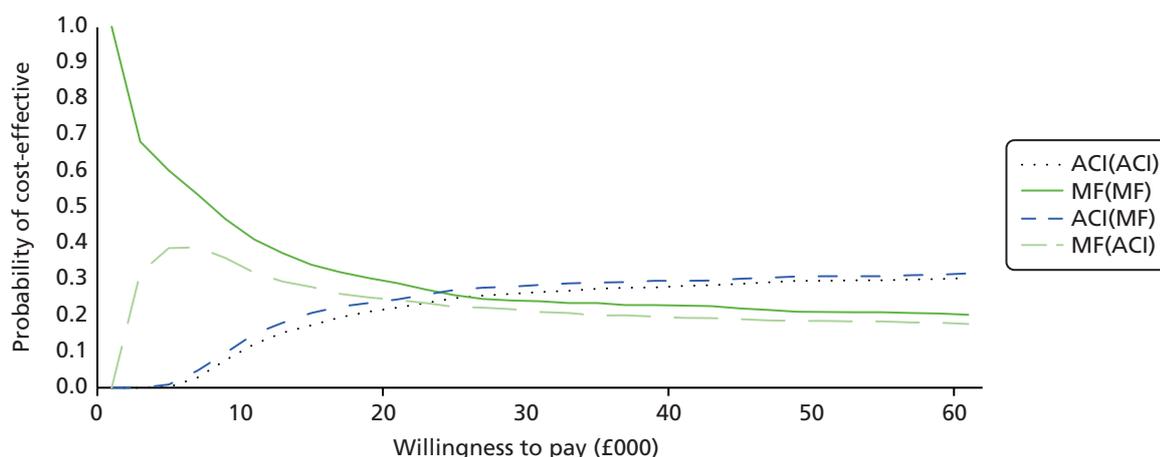


FIGURE 30 Cost-effectiveness acceptability curves: MF as a day case. (a) MF as a day case – scenario 1; and (b) MF as a day case – scenario 2.

TABLE 31 Sensitivity cost-effectiveness results: MF procedure as a day case surgery and not as an inpatient

| Procedure | Total mean costs (£) | Total mean QALYs | Comparison | Incremental costs (£) | Incremental QALYs | ICER (£) (cost per QALY gained) |
|----------------------|----------------------|------------------|----------------------|-----------------------|---------------------|---------------------------------|
| Deterministic | | | | | | |
| MF(MF) | 2819 | 17.0033 | – | – | – | – |
| MF(ACI) | 4621 | 17.0284 | MF(ACI) vs. MF(MF) | 1802 | 0.0251 ^a | Extendedly dominated |
| ACI(MF) | 19,756 | 17.9570 | ACI(MF) vs. MF(MF) | 16,937 | 0.9537 | 17,758 |
| ACI(ACI) | 20,921 | 18.0228 | ACI(ACI) vs. ACI(MF) | 1165 | 0.0658 ^a | 17,715 |
| Probabilistic | | | | | | |
| MF(MF) | 2811 | 17.0137 | – | – | – | – |
| MF(ACI) | 4620 | 17.0412 | MF(ACI) vs. MF(MF) | 1809 | 0.0275 ^a | Extendedly dominated |
| ACI(MF) | 19,788 | 17.9065 | ACI(MF) vs. MF(MF) | 16,977 | 0.8928 | 19,017 |
| ACI(ACI) | 20,951 | 17.9975 | ACI(ACI) vs. ACI(MF) | 1163 | 0.0910 ^a | 12,777 |

^a As the incremental QALYs are near zero, the ICER can fluctuate widely.

**FIGURE 31** Cost-effectiveness acceptability curve: MF as a day case procedure.

Sensitivity analysis 4: improving the success rates of microfracture

In this SA we have conducted a 'what if' scenario, for which we have assumed that the duration of success for MF increases (1) by 20% and (2) by 40%. Table 32 shows the sensitivity cost-effectiveness results by scenario for the increase in the duration of success for MF. ACI was still more costly than MF and there were more QALYs gained with ACI than MF; even though there was a slight fall in the incremental QALYs gained compared with the base-case results, this was not enough to change the ICERs. For both scenarios, ACI as a first repair was more cost-effective than MF as a first repair. These results were of similar magnitudes and directions for the probabilistic results.

Figure 32a–d presents the CEACs for the SA for the increase in the duration of success for MF for scenarios 1 and 2. For a 20% increase in the duration of success for MF: scenario 1, if the decision-maker was willing to pay £20,000, the probability that ACI was more cost-effective than MF was 58% (see Figure 32a); scenario 2,

TABLE 32 Sensitivity cost-effectiveness results (by scenario): improving the success rates of MF

| Procedure | Total mean costs (£) | Total mean QALYs | Incremental costs (£) | Incremental QALYs | ICER (£) (cost per QALY gained) |
|---|----------------------|------------------|-----------------------|-------------------|---------------------------------|
| Deterministic: 20% increase in success rates | | | | | |
| <i>Scenario 1</i> | | | | | |
| MF(ACI) | 6392 | 17.0756 | – | – | – |
| ACI(ACI) | 20,921 | 18.0228 | 14,529 | 0.9472 | 15,338 |
| <i>Scenario 2</i> | | | | | |
| MF(MF) | 4969 | 17.0607 | – | – | – |
| ACI(MF) | 19,892 | 17.9639 | 14,923 | 0.9033 | 16,521 |
| Probabilistic: 20% increase in success rates | | | | | |
| <i>Scenario 1</i> | | | | | |
| MF(ACI) | 6387 | 17.0546 | – | – | – |
| ACI(ACI) | 20,895 | 18.0271 | 14,509 | 0.9725 | 14,919 |
| <i>Scenario 2</i> | | | | | |
| MF(MF) | 4954 | 17.0403 | – | – | – |
| ACI(MF) | 19,856 | 17.9429 | 14,901 | 0.9027 | 16,508 |
| Deterministic: 40% increase in success rates | | | | | |
| <i>Scenario 1</i> | | | | | |
| MF(ACI) | 5698 | 17.0787 | – | – | – |
| ACI(ACI) | 20,921 | 18.0228 | 15,223 | 0.9441 | 16,125 |
| <i>Scenario 2</i> | | | | | |
| MF(MF) | 4820 | 17.0758 | – | – | – |
| ACI(MF) | 19,892 | 17.9741 | 15,072 | 0.8983 | 16,778 |
| Probabilistic: 40% increase in success rates | | | | | |
| <i>Scenario 1</i> | | | | | |
| MF(ACI) | 5682 | 17.0535 | – | – | – |
| ACI(ACI) | 20,975 | 18.0331 | 15,292 | 0.9796 | 15,611 |
| <i>Scenario 2</i> | | | | | |
| MF(MF) | 4803 | 17.0520 | – | – | – |
| ACI(MF) | 19,939 | 17.9887 | 15,136 | 0.9367 | 16,159 |

the probability that ACI was more cost-effective than MF was 56% (see Figure 32b). For a 40% increase in the duration of success for MF these probability figures were 57% (see Figure 32c) and 55% (see Figure 32d), respectively.

Table 33 presents the deterministic and probabilistic cost-effectiveness results. The costs of MF have fallen slightly and also the QALYs for MF have increased. ACI(ACI) has an ICER of just under £18,000 (for a 20% increase in the duration of success rate for MF) and over £21,000 (for a 40% increase in the duration of success rate for MF).

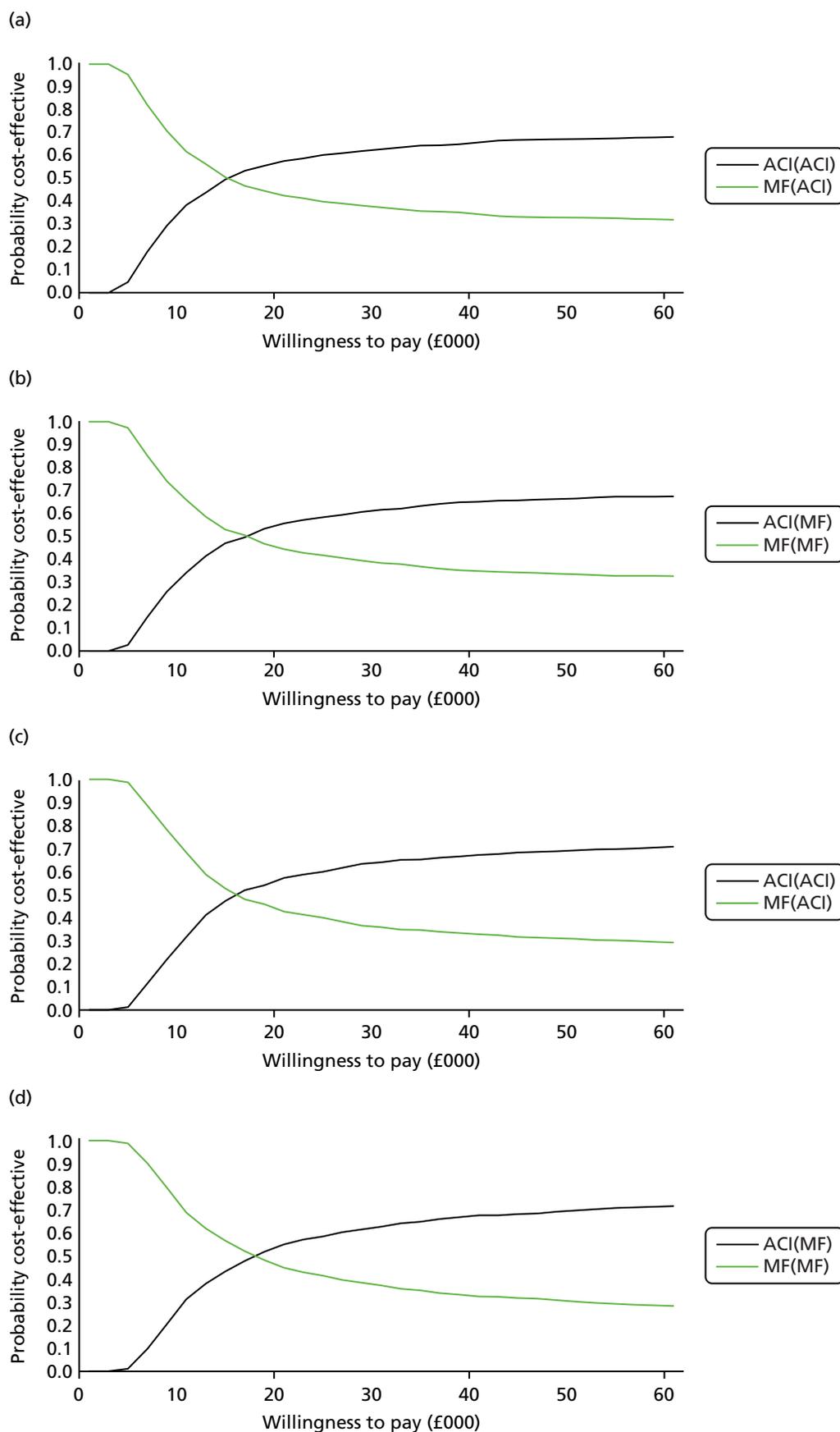


FIGURE 32 Cost-effectiveness acceptability curves: increase in MF success rate: (a) 20% increase in MF success rate – scenario 1; (b) 20% increase in MF success rate – scenario 2; (c) 40% increase in MF success rate – scenario 1; and (d) 40% increase in MF success rate – scenario 2.

TABLE 33 Sensitivity cost-effectiveness results: improving the success rates of MF

| Procedure | Total mean costs (£) | Total mean QALYs | Comparison | Incremental costs (£) | Incremental QALYs | ICER (£) (cost per QALY gained) |
|---|----------------------|------------------|----------------------|-----------------------|---------------------|---------------------------------|
| Deterministic: 20% increase in success rates | | | | | | |
| MF(MF) | 4969 | 17.0607 | – | – | – | – |
| MF(ACI) | 6392 | 17.0756 | MF(ACI) vs. MF(MF) | 1423 | 0.0149 ^a | Extendedly dominated |
| ACI(MF) | 19,892 | 17.9639 | ACI(MF) vs. MF(MF) | 14,923 | 0.9033 | 16,521 |
| ACI(ACI) | 20,921 | 18.0228 | ACI(ACI) vs. ACI(MF) | 1029 | 0.0589 ^a | 17,480 |
| Probabilistic: 20% increase in success rates | | | | | | |
| MF(MF) | 4954 | 17.0403 | – | – | – | – |
| MF(ACI) | 6387 | 17.0546 | MF(ACI) vs. MF(MF) | 1433 | 0.0144 ^a | Extendedly dominated |
| ACI(MF) | 19,856 | 17.9429 | ACI(MF) vs. MF(MF) | 14,901 | 0.9027 | 16,508 |
| ACI(ACI) | 20,895 | 18.0271 | ACI(ACI) vs. ACI(MF) | 1040 | 0.0842 ^a | 12,356 |
| Deterministic: 40% increase in success rates | | | | | | |
| MF(MF) | 4820 | 17.0758 | – | – | – | – |
| MF(ACI) | 5698 | 17.0787 | MF(ACI) vs. MF(MF) | 878 | 0.0029 ^a | Extendedly dominated |
| ACI(MF) | 19,892 | 17.9741 | ACI(MF) vs. MF(MF) | 15,072 | 0.8983 | 16,778 |
| ACI(ACI) | 20,921 | 18.0228 | ACI(ACI) vs. ACI(MF) | 1029 | 0.0487 ^a | 21,130 |
| Probabilistic: 40% increase in success rates | | | | | | |
| MF(MF) | 4803 | 17.0520 | – | – | – | – |
| MF(ACI) | 5682 | 17.0535 | MF(ACI) vs. MF(MF) | 880 | 0.0015 ^a | Extendedly dominated |
| ACI(MF) | 19,939 | 17.9887 | ACI(MF) vs. MF(MF) | 15,136 | 0.9367 | 16,159 |
| ACI(ACI) | 20,975 | 18.0331 | ACI(ACI) vs. ACI(MF) | 1036 | 0.0443 ^a | 23,356 |

^a As the incremental QALYs are near zero, the ICER can fluctuate widely.

Figure 33a and b presents the CEACs for the SA results for the ‘what if’ scenario if the duration of success of MF was to increase by 20% and 40%, respectively. For a 20% increase in the MF success rate, the graph suggests that if the decision-maker is willing to pay only £20,000 then there is not much difference in the four options; however, if the decision-maker is willing to pay £22,000 or more then ACI as a first repair is more cost-effective than MF as a first repair. For a 40% increase in the MF success rate, the graph indicates that if the decision-maker is willing to pay £20,000 then ACI as a first repair [ACI(ACI) or ACI(MF)] is more cost-effective than MF (approximately 32–33% probability that it is more cost-effective).

Sensitivity analysis 5: starting age of cohort is 45 years

In the base-case analysis the starting age for the cohort was 33 years. In this SA the starting age is changed to 45 years (patients are nearer to the knee replacement age) to see how this affects the ICER. Table 34 presents the deterministic and probabilistic cost-effectiveness results by scenario. Even though the model is starting at a later age (45 years), the results are very similar to the base case, with ACI as a first repair being cost-effective compared with MF as a first repair.

Figure 34a and b presents the CEACs for the SA with a starting age of 45 years for the cohort for scenarios 1 and 2. For scenarios 1 and 2, if the decision-maker was willing to pay £20,000, the probability that ACI was cost-effective relative to MF was 64%.

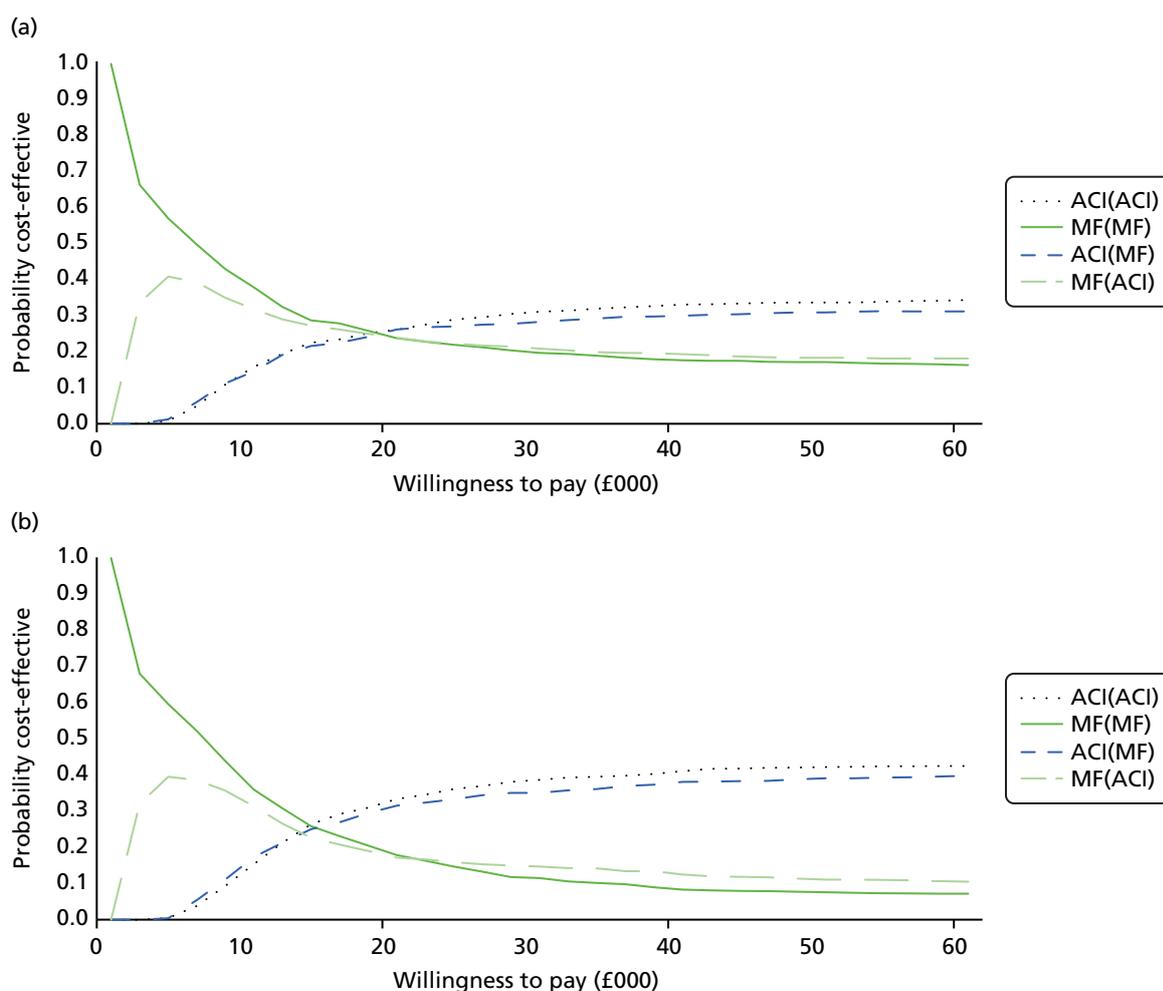


FIGURE 33 Cost-effectiveness acceptability curves: increase in MF success rate. (a) Increase in MF success rate by 20%; and (b) increase in MF success rate by 40%. Note that MF only becomes cost-effective if the duration of benefit is much longer.

TABLE 34 Sensitivity cost-effectiveness results (by scenario): starting age for cohort is 45 years

| Procedure | Total mean costs (£) | Total mean QALYs | Incremental costs (£) | Incremental QALYs | ICER (£) (cost per QALY gained) |
|----------------------|----------------------|------------------|-----------------------|-------------------|---------------------------------|
| Deterministic | | | | | |
| <i>Scenario 1</i> | | | | | |
| MF(ACI) | 6422 | 15.0445 | – | – | – |
| ACI(ACI) | 16,784 | 16.0327 | 10,362 | 0.9882 | 10,486 |
| <i>Scenario 2</i> | | | | | |
| MF(MF) | 5267 | 15.0187 | – | – | – |
| ACI(MF) | 16,116 | 15.9766 | 10,849 | 0.9579 | 11,326 |
| Probabilistic | | | | | |
| <i>Scenario 1</i> | | | | | |
| MF(ACI) | 6441 | 15.0833 | – | – | – |
| ACI(ACI) | 16,724 | 16.0377 | 10,283 | 0.9544 | 10,755 |
| <i>Scenario 2</i> | | | | | |
| MF(MF) | 5281 | 14.9900 | – | – | – |
| ACI(MF) | 16,053 | 15.9562 | 10,772 | 0.9962 | 11,149 |

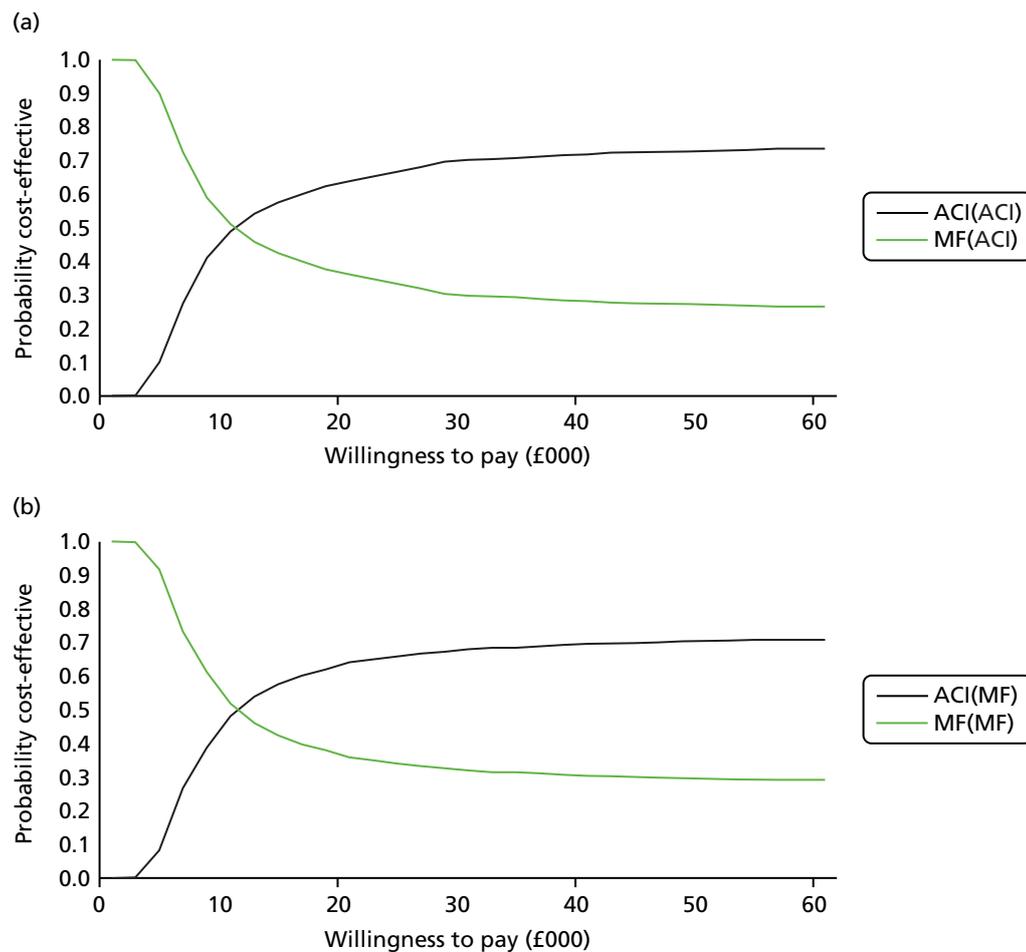


FIGURE 34 Cost-effectiveness acceptability curves: starting age is 45 years: (a) scenario 1; and (b) scenario 2.

Table 35 shows the deterministic and probabilistic cost-effectiveness results with a starting age of 45 years for the cohort and results were ranked by the least costly option. Results were similar to the base-case results and ACI(ACI) remained the most cost-effective procedure.

Figure 35 presents the CEAC for a starting age of 45 years for the cohort. If the decision-maker is willing to pay £20,000 per QALY then either ACI(ACI) or ACI(MF) are the most cost-effective options.

Discussion

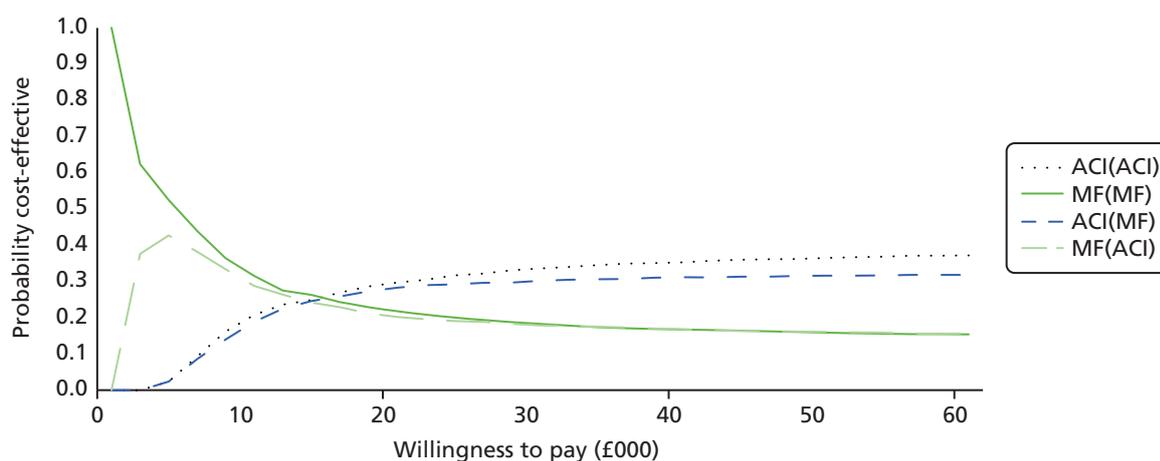
For the base-case analysis, a hypothetical cohort of 1000 patients with cartilage knee defects with a starting age of 33 years was followed over a lifetime horizon. The cycle length for the model was set to 1 year. The analysis was conducted from the perspective of the NHS and PSS. Data for the transition probabilities, mortality rates and utilities were obtained from the literature. Health outcomes were measured in QALYs. The majority of unit costs were obtained from the NHS reference costs¹⁵¹ database and all costs are in pounds sterling in 2012–13 prices. Results were compared in two different ways for ACI and MF:

- First, we used two scenarios: scenario 1 with all second repairs ACI and scenario 2 with all second repairs MF.
- Second, all four options were ranked in order of increasing costs; any options (sequences) that were more expensive and less effective were excluded (simple dominance); and if the ICER for a given category was higher than that of the next more effective alternative then this category was eliminated (extended dominance).

TABLE 35 Sensitivity cost-effectiveness results: starting age for cohort is 45 years

| Procedure | Total mean costs (£) | Total mean QALYs | Comparison | Incremental costs (£) | Incremental QALYs | ICER (£) (cost per QALY gained) |
|----------------------|----------------------|------------------|----------------------|-----------------------|---------------------|---------------------------------|
| Deterministic | | | | | | |
| MF(MF) | 5267 | 15.0187 | – | – | – | – |
| MF(ACI) | 6422 | 15.0445 | MF(ACI) vs. MF(MF) | 1155 | 0.0258 ^a | Extendedly dominated |
| ACI(MF) | 16,116 | 15.9766 | ACI(MF) vs. MF(MF) | 10,849 | 0.9579 | 11,326 |
| ACI(ACI) | 16,784 | 16.0327 | ACI(ACI) vs. ACI(MF) | 667 | 0.0561 ^a | 11,898 |
| Probabilistic | | | | | | |
| MF(MF) | 5281 | 14.9900 | – | – | – | – |
| MF(ACI) | 6441 | 15.0833 | MF(ACI) vs. MF(MF) | 1160 | 0.0933 ^a | Extendedly dominated |
| ACI(MF) | 16,053 | 15.9562 | ACI(MF) vs. MF(MF) | 10,772 | 0.9662 | 11,149 |
| ACI(ACI) | 16,724 | 16.0377 | ACI(ACI) vs. ACI(MF) | 671 | 0.0815 ^a | 8241 |

^a As the incremental QALYs are near zero, the ICER can fluctuate widely.

**FIGURE 35** Cost-effectiveness acceptability curve: starting age of cohort is 45 years.

Results are expressed as incremental cost per QALY gained. An annual discount rate of 3.5% was applied to both costs and QALYs. We ran the model deterministically and probabilistically with 1000 iterations. We undertook various SAs. These bootstrapped iterations were used to construct the CEACs. The CEACs were presented using a willingness-to-pay threshold from £0 to £60,000.

Methods and summary of findings

For the base-case analysis, for the discounted deterministic results MF was the less costly option but had fewer QALYs, whereas ACI was the more expensive option but generated more QALYs.

For scenario 1, the cost per QALY gained for ACI compared with MF was £14,395, and for scenario 2 the cost per QALY gained £15,598. These results were confirmed by the CEACs: so if the decision-maker is willing to pay £20,000 for a QALY, ACI is 56–59% more likely to be cost-effective than MF. For both scenarios, ACI as a first repair appeared more cost-effective than MF as a first repair.

When looking at the different sequences (options), the initial ACI appears more cost-effective than initial MF, and, for those who need a second repair after the first ACI, this should also be another ACI.

The option MF(ACI) was extendedly dominated by a linear combination of MF(MF) and ACI(MF) and therefore this option was eliminated from the comparison. For the different sequences, the CEAC for the base-case results confirmed these results and showed that if the decision-maker is willing to pay £18,000 or more for a QALY than ACI as a first procedure is more cost-effective than MF as a first procedure.

We found that the key cost driver was the cost of the cells for the ACI procedure, but, over the time horizon, ACI is more beneficial (more gain in QALYs) and cost saving to the NHS (fewer people in need of a second repair or of a TKR).

A number of SAs were undertaken to determine the cost-effectiveness of various options and the majority of results were in line with the base-case analysis. However, we found that the model was sensitive to the cost of cells – we know that these are not the true costs as the NHS receives confidential discounts from the manufacturers. This means that with the cell cost reduction ACI(ACI) is likely to be even more cost-effective than the base-case cost per QALY ratio that was presented. We also found that the model was sensitive to the time horizon: with a shorter time horizon – 10 years – the cost per QALY for the two initial ACI options rose to around £26,000 because of the costs of the ACI procedure occurring at the start and the benefits of ACI not being realised until much later, such as the reduced need for TKRs. When the time horizon was longer, the model results were in line with the base-case results. The SAs conducted using Oswestry data found ACI not to be cost-effective compared with MF and mainly due to the lower utility value in the fourth year for ACI compared with MF. However, for reasons explained in *Chapter 5*, there are several confounding factors that influence these utility values.

Strengths and limitations

If the first repair fails, the Markov model considers patients having a maximum of two knee repairs (any combination of MF and ACI) if they choose to.

However, the model does have a number of limitations. First, the length of follow-up we found in the trials published in the literature was too short and hence, there are no long-term data on the success and failure rates (including long-term benefits and AEs) for each of these procedures and what the average age is for these patients when a TKR/PKR is required. Hence, we relied heavily on two papers^{70,100} to obtain transition probabilities for our economic model. In addition, because of the paucity of data from clinical studies, transition probabilities were not available for each transition in the model; hence we made a number of assumptions for our model, which were checked for plausibility with our clinical advisors. However, results from the long-term ACTIVE trial³⁵ (comparing ACI/MACI with standard treatments) and the TOPKAT trial (comparing TKR with PKR)¹⁶⁴ will provide useful information with which to populate our economic model, although results will not be available until 2017 and 2019, respectively.

Second, because of the short follow-up, we also found that there were no long-term data on utility values that were associated with each of these procedures. We have had to rely heavily on the literature and on a few studies in particular, such as Gerlier *et al.*¹⁴¹ Also, we found no studies that mapped any of the clinical measures – such as Lysholm score or the KOOS – to either the EQ-5D or SF-6D to generate utility values, which would have been helpful in our model.

Third, we relied on clinical experience for information on the average resources used (e.g. outpatient and rehabilitation visits) for these patients.

Fourth, we did not take into account any costs for the analgesics, based on advice from our clinical experts, as these costs are negligible and would not have altered the base-case cost per QALY. Also, not all of the costs obtained were from the NHS reference costs. We used the previous HTA report by Clar *et al.*,³ who obtained costs from Aberdeen and Southampton hospitals to populate their economic model. Although these costs were inflated to 2012–13 prices using the HCHS index,¹⁵⁴ to get a more accurate picture of these costs it would have been better to have carried out 'bottom-up costing'.

Fifth, the model has not taken into account any private patient costs, such as time off work and loss of pay (productivity): this population, of those who have either an ACI or MF, is a young cohort and it will primarily have an effect on their own costs. In line with this, it would have been interesting to know how long it would take this population cohort to return to normal activities after each of these procedures (return to work or sports).

Finally, we did not include any AEs, as there were no key differences between the two treatment arms.

Additional analyses based on survival analysis

Next, we report the results of the additional economic analyses undertaken, incorporating new parameter values, in particular the survival curves for failure rates reported in *Chapter 3*. We have used data from the long-term studies in a new base-case analysis, using the whole Nawaz cohort⁸⁰ results for ACI, and pooling the MF results from the three studies. At the request of NICE, we have used an implantation cost of £2396 and also omitted the option for MF failure followed by another MF. Unless specified, the model structure and parameter values remain the same as those in the initial report.

The different sequences of procedures were ranked in order of increasing cost. We eliminated any categories for which another category was cheaper and more effective (simple dominance). If the ICER for a given category is higher than that of the next more effective alternative this category was eliminated (extended dominance). For the remaining options, we reported the ICERs, measured as cost per QALY gained.

When QALY differences are small, the probabilistic ICERs will fluctuate quite a lot. The deterministic ICERs are more reliable.

New base case

- Data used for ACI failure rates: Nawaz *et al.*⁸⁰ (whole cohort).
- Data used for MF failure rates: pooled data (Layton 2015,¹¹⁴ Knutsen 2007,⁶⁷ Saris 2009⁷⁰).
- Cost of harvesting: £870.
- Cost of implantation: £2396, as requested by NICE. This includes an inpatient stay. ACI can be done on a day case basis, though it should be noted that, because it is often provided as a specialist 'regional' service, overnight stays may be unavoidable because of distance. The clinical authors of this report vary between one-night stays for all and some being done as day case. The operation is often open, and such exposure is much more painful than the arthroscopic surgery used for harvesting the initial tissue). However, mini-arthrotomy may be used.
- MF is nearly always a day case procedure.

Table 36 presents the results from the new base-case analysis. The discounted deterministic results show that MF(ACI) was the least costly option and had the fewest QALYs; although ACI(ACI) generated the most QALYs, it was also the most expensive option. The option of ACI(MF) was extendedly dominated by a linear combination of MF(ACI) and ACI(ACI), and therefore this option was eliminated from the comparison. The ICER comparing ACI(ACI) with MF(ACI) was just under £19,000; doing ACI first is more cost-effective. The discounted probabilistic results were very similar. *Figure 36* shows the cost-effectiveness analysis for the two remaining options. The graph shows that, for amounts below £20,000, MF(ACI) is the most cost-effective option; at a willingness to pay of £20,000 there is not much difference between the two options, and, at a willingness to pay above £20,000, ACI(ACI) is probably more cost-effective.

TABLE 36 Base-case deterministic and probabilistic cost-effectiveness results

| Procedure | Total mean costs (£) | Total mean QALYs | Incremental costs (£) | Incremental QALYs | ICER vs. MF(ACI) | Comparator |
|------------------------------------|----------------------|------------------|-----------------------|-------------------|--------------------|------------|
| Deterministic: undiscounted | | | | | | |
| MF(ACI) | 7736 | 34.2885 | – | – | – | – |
| ACI(MF) | 22,661 | 35.5596 | 14,926 | 1.2711 | Extended dominated | MF(ACI) |
| ACI(ACI) | 24,134 | 35.6999 | 1473 | 0.1403 | 11,619 | MF(ACI) |
| Deterministic: discounted | | | | | | |
| MF(ACI) | 6248 | 17.1350 | – | – | – | – |
| ACI(MF) | 21,400 | 17.9304 | 15,152 | 0.7954 | Extended dominated | MF(ACI) |
| ACI(ACI) | 22,461 | 17.9953 | 1062 | 0.0650 | 18,844 | MF(ACI) |
| Probabilistic: discounted | | | | | | |
| MF(ACI) | 6261 | 17.1523 | – | – | – | – |
| ACI(MF) | 21,410 | 17.9048 | 15,210 | 0.7525 | Extended dominated | MF(ACI) |
| ACI(ACI) | 22,532 | 17.9872 | 1061 | 0.0824 | 19,487 | MF(ACI) |

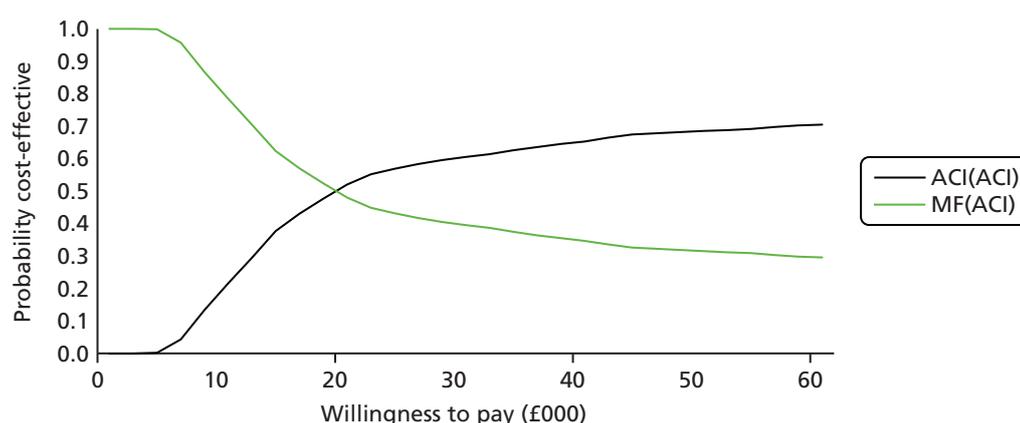


FIGURE 36 Cost-effectiveness acceptability curve: base case.

Sensitivity analyses on prices

Note that in analysis of price changes, the QALY gain does not change in the deterministic arms, as expected. However, when the model is run probabilistically all of the input variables change because of the different distributions, hence both the costs and QALYs will change.

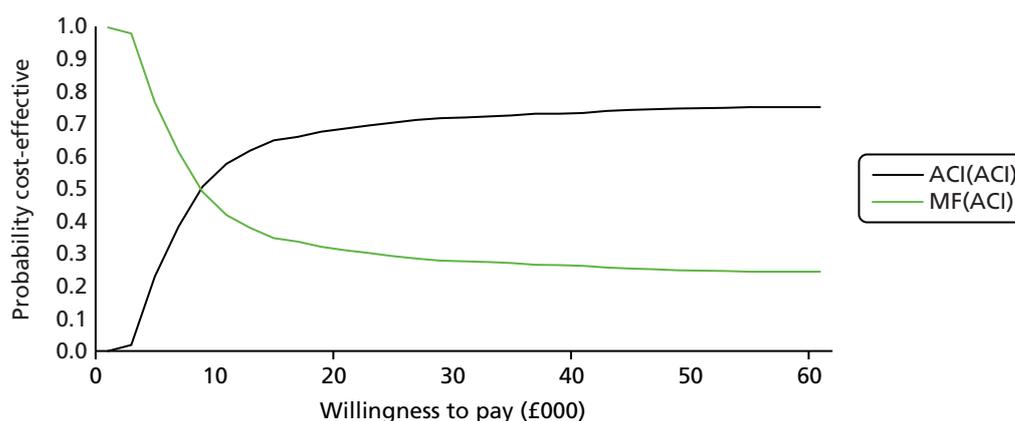
Cells at £6000

- Data used for ACI failure rates: Nawaz *et al.*⁸⁰ (whole cohort).
- Data used for MF failure rates: pooled data (Layton 2015,¹¹⁴ Knutsen 2007,⁶⁷ Saris 2009⁷⁰).
- Cost of harvesting: £870.
- Cost of implantation: £2396.
- Cost of cells: £6000.

Table 37 shows the results when the price of cells is reduced to £6000. The deterministic discounted results highlighted that the option of ACI(MF) was extendedly dominated and therefore the ICER comparing ACI(ACI) with MF(ACI) was just under £7500. Figure 37 shows that at a willingness to pay of £8000 there is not much difference between the two remaining options.

TABLE 37 Deterministic and probabilistic cost-effectiveness results: £6000 price

| Procedure | Total mean costs (£) | Total mean QALYs | Incremental costs (£) | Incremental QALYs | ICER vs. MF(ACI) | Comparator |
|------------------------------------|----------------------|------------------|-----------------------|-------------------|--------------------|------------|
| Deterministic: undiscounted | | | | | | |
| MF(ACI) | 6771 | 34.2885 | – | – | – | – |
| ACI(MF) | 12,661 | 35.5596 | 5,890 | 1.2711 | Extended dominated | MF(ACI) |
| ACI(ACI) | 13,244 | 35.6999 | 583 | 0.1403 | 4586 | MF(ACI) |
| Deterministic: discounted | | | | | | |
| MF(ACI) | 5441 | 17.1350 | – | – | – | – |
| ACI(MF) | 11,400 | 17.9304 | 5,959 | 0.7954 | Extended dominated | MF(ACI) |
| ACI(ACI) | 11,820 | 17.9953 | 420 | 0.0650 | 7414 | MF(ACI) |
| Probabilistic: discounted | | | | | | |
| MF(ACI) | 5452 | 17.1340 | – | – | – | – |
| ACI(MF) | 11,486 | 17.9110 | 6,034 | 0.7770 | 7766 | MF(ACI) |
| ACI(ACI) | 11,909 | 17.9474 | 423 | 0.0364 | 11,622 | ACI(MF) |

**FIGURE 37** Cost-effectiveness acceptability curve: £6000.

Cells at cost £8000

- Data used for ACI failure rates: Nawaz *et al.*⁸⁰ (whole cohort).
- Data used for MF failure rates: pooled data (Layton 2015,¹¹⁴ Knutsen 2007,⁶⁷ Saris 2009⁷⁰).
- Cost of harvesting: £870.
- Cost of implantation: £2396.
- Cost of cells: £8000.

Table 38 shows the results when the price of cells is reduced to £8000. The deterministic discounted results highlighted that the option of ACI(MF) was extendedly dominated and therefore the ICER comparing ACI(ACI) with MF(ACI) was just under £10,000. Figure 38 shows that at a willingness to pay of £10,000 there is not much difference between the two remaining options.

TABLE 38 Deterministic and probabilistic cost-effectiveness results: £8000 price

| Procedure | Total mean costs (£) | Total mean QALYs | Incremental costs (£) | Incremental QALYs | ICER vs. MF(ACI) | Comparator |
|------------------------------------|----------------------|------------------|-----------------------|-------------------|--------------------|------------|
| Deterministic: undiscounted | | | | | | |
| MF(ACI) | 6964 | 34.2885 | – | – | – | – |
| ACI(MF) | 14,661 | 35.5596 | 7697 | 1.2711 | Extended dominated | MF(ACI) |
| ACI(ACI) | 15,422 | 35.6999 | 761 | 0.1403 | 5993 | MF(ACI) |
| Deterministic: discounted | | | | | | |
| MF(ACI) | 5602 | 17.1350 | – | – | – | – |
| ACI(MF) | 13,400 | 17.9304 | 7797 | 0.7954 | Extended dominated | MF(ACI) |
| ACI(ACI) | 13,948 | 17.9953 | 549 | 0.0650 | 9700 | MF(ACI) |
| Probabilistic: discounted | | | | | | |
| MF(ACI) | 5608 | 17.1630 | – | – | – | – |
| ACI(MF) | 13,430 | 17.9500 | 7822 | 0.7871 | 9938 | MF(ACI) |
| ACI(ACI) | 13,983 | 18.0242 | 553 | 0.0742 | 7454 | ACI(MF) |

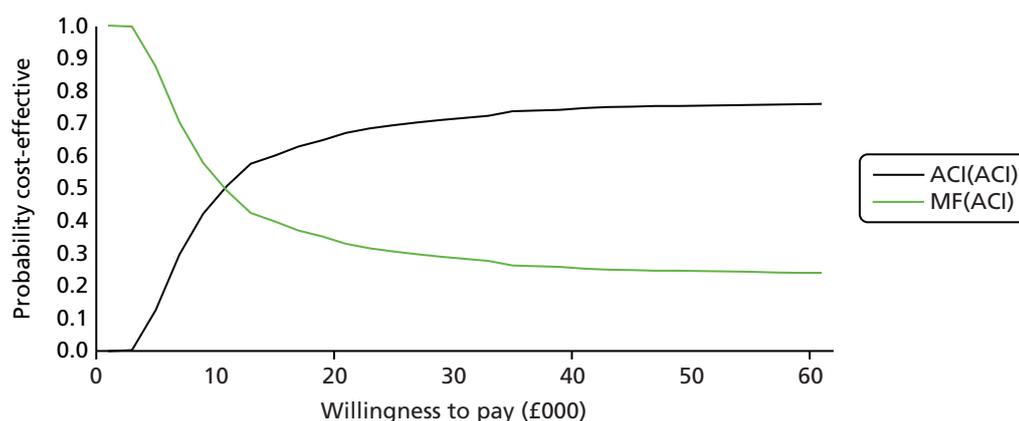


FIGURE 38 Cost-effectiveness acceptability curve: £8000.

Cells at cost of £12,000

- Data used for ACI failure rates: Nawaz *et al.*⁸⁰ (whole cohort).
- Data used for MF failure rates: pooled data (Layton 2015,¹¹⁴ Knutsen 2007,⁶⁷ Saris 2009⁷⁰).
- Cost of harvesting: £870.
- Cost of implantation: £2396.
- Cost of cells: £12,000.

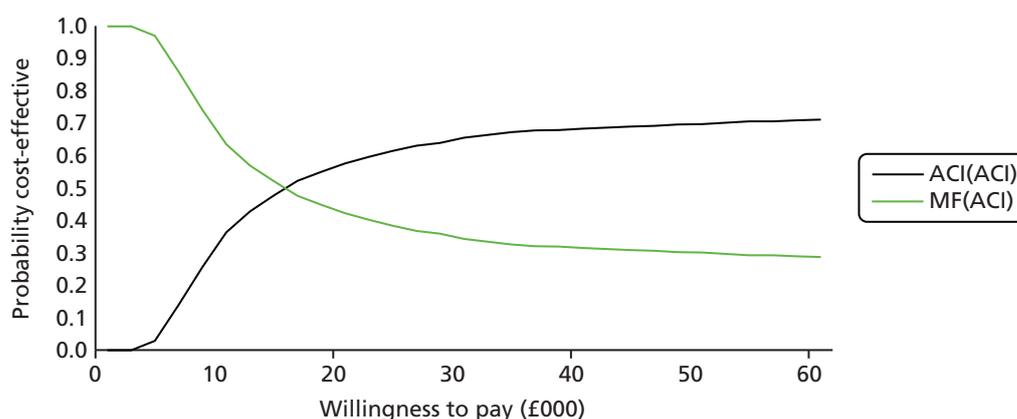
Table 39 shows the results when the price of cells is reduced to £12,000. The deterministic discounted results highlighted that the option of ACI(MF) was extendedly dominated and therefore the ICER comparing ACI(ACI) with MF(ACI) was just under £14,500. Figure 39 shows that at a willingness to pay of £16,000 there is not much difference between the two remaining options.

Sensitivity analyses: post-repair utility

In our first assessment report, we assumed that patients who decided not to have a further repair had had some benefit, and had improved from a utility of 0.654 before the repair to 0.691 afterwards. NICE asked

TABLE 39 Deterministic and probabilistic cost-effectiveness results: £12,000 price

| Procedure | Total mean costs (£) | Total mean QALYs | Incremental costs (£) | Incremental QALYs | ICER vs. MF(ACI) | Comparator |
|------------------------------------|----------------------|------------------|-----------------------|-------------------|--------------------|------------|
| Deterministic: undiscounted | | | | | | |
| MF(ACI) | 7350 | 34.2885 | – | – | – | – |
| ACI(MF) | 18,661 | 35.5596 | 11,312 | 1.2711 | Extended dominated | MF(ACI) |
| ACI(ACI) | 19,778 | 35.6999 | 1117 | 0.1403 | 8806 | MF(ACI) |
| Deterministic: discounted | | | | | | |
| MF(ACI) | 5925 | 17.1350 | – | – | – | – |
| ACI(MF) | 17,400 | 17.9304 | 11,475 | 0.7954 | Extended dominated | MF(ACI) |
| ACI(ACI) | 18,205 | 17.9953 | 805 | 0.0650 | 14,272 | MF(ACI) |
| Probabilistic: discounted | | | | | | |
| MF(ACI) | 5918 | 17.1425 | – | – | – | – |
| ACI(MF) | 17,320 | 17.9539 | 11,402 | 0.8114 | Extended dominated | MF(ACI) |
| ACI(ACI) | 18,131 | 17.9899 | 811 | 0.0360 | 14,412 | MF(ACI) |

**FIGURE 39** Cost-effectiveness acceptability curve: £12,000.

us to assess the effect of several assumptions for utilities in those in whom repair is unsuccessful, but who choose not to have another operation.

Utility for choose no second repair set to same as failure

- Data used for ACI failure rates: Nawaz *et al.*⁸⁰ (whole cohort).
- Data used for MF failure rates: pooled data (Layton 2015,¹¹⁴ Knutsen 2007,⁶⁷ Saris 2009⁷⁰).
- Cost of harvesting: £870.
- Cost of implantation: £2396.
- Cost of cells: £16,000.

Utility for those who choose no second repair: 0.654.

Table 40 shows the results when the utility value for those who choose no second repair is 0.654. The deterministic discounted results highlighted that the option of ACI(MF) was extendedly dominated and therefore the ICER comparing ACI(ACI) with MF(ACI) was just over £15,500 (results are similar to the new base-case results). Figure 40 shows that at a willingness to pay of £16,000 there is not much difference between the two remaining options.

TABLE 40 Deterministic and probabilistic cost-effectiveness results: utility = 0.654

| Procedure | Total mean costs (£) | Total mean QALYs | Incremental costs (£) | Incremental QALYs | ICER vs. MF(ACI) | Comparator |
|------------------------------------|----------------------|------------------|-----------------------|-------------------|--------------------|------------|
| Deterministic: undiscounted | | | | | | |
| MF(ACI) | 7736 | 32.8665 | – | – | – | – |
| ACI(MF) | 22,661 | 34.4351 | 14,926 | 1.5686 | Extended dominated | MF(ACI) |
| ACI(ACI) | 24,134 | 34.6021 | 1473 | 0.1670 | 9449 | MF(ACI) |
| Deterministic: discounted | | | | | | |
| MF(ACI) | 6248 | 16.5058 | – | – | – | – |
| ACI(MF) | 21,400 | 17.4667 | 15,152 | 0.9609 | Extended dominated | MF(ACI) |
| ACI(ACI) | 22,461 | 17.5428 | 1062 | 0.0762 | 15,634 | MF(ACI) |
| Probabilistic: discounted | | | | | | |
| MF(ACI) | 6247 | 16.4820 | – | – | – | – |
| ACI(MF) | 21,446 | 17.4583 | 15,198 | 0.9763 | Extended dominated | MF(ACI) |
| ACI(ACI) | 22,2522 | 17.5499 | 1077 | 0.0916 | 15,241 | MF(ACI) |

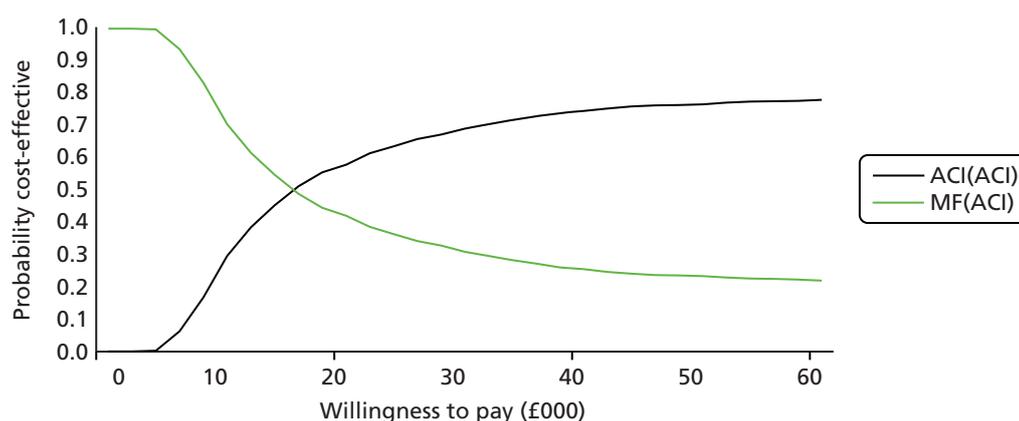


FIGURE 40 Cost-effectiveness acceptability curve: utility = 0.654.

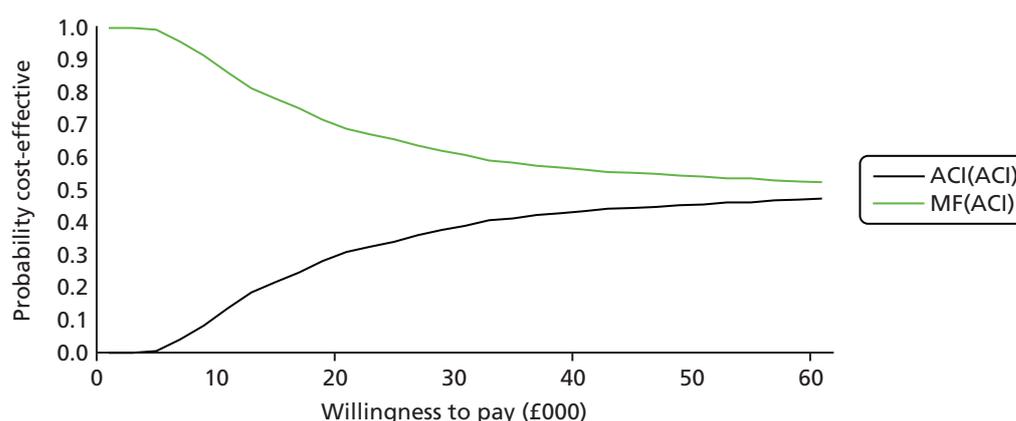
Utility for choose no second repair set to same as success

- Data used for ACI failure rates: Nawaz *et al.*⁸⁰ (whole cohort).
- Data used for MF failure rates: pooled data (Layton 2015,¹¹⁴ Knutsen 2007,⁶⁷ Saris 2009⁷⁰).
- Cost of harvesting: £870.
- Cost of implantation: £2396.
- Cost of cells: £16,000.
- Utility for those who choose no second repair: 0.817. Note that this assumption greatly increases utility gain among those who do not get good results after MF, and reduces the marginal QALY gains from ACI.

Table 41 shows the results when the utility value for those who choose no second repair is 0.817. Here we are assuming that there are increases in utility gain among those who do not get good results after MF, and therefore this reduces the marginal QALY gains from ACI. The deterministic discounted results highlighted that the option of ACI(MF) was extendedly dominated and therefore the ICER comparing ACI(ACI) with MF(ACI) was just over £62,000. Hence, the option ACI(ACI) is no longer considered to be cost-effective. Figure 41 shows that, at a willingness to pay of £20,000, MF(ACI) is 70% more likely to be cost-effective than ACI(ACI).

TABLE 41 Deterministic and probabilistic cost-effectiveness results: utility = 0.817

| Procedure | Total mean costs (£) | Total mean QALYs | Incremental costs (£) | Incremental QALYs | ICER vs. MF(ACI) | Comparator |
|------------------------------------|----------------------|------------------|-----------------------|-------------------|--------------------|------------|
| Deterministic: undiscounted | | | | | | |
| MF(ACI) | 7736 | 39.1309 | – | – | – | – |
| ACI(MF) | 22,661 | 39.3889 | 14,926 | 0.2580 | Extended dominated | MF(ACI) |
| ACI(ACI) | 24,134 | 39.4383 | 1473 | 0.0494 | 53,352 | MF(ACI) |
| Deterministic: discounted | | | | | | |
| MF(ACI) | 6248 | 19.2776 | – | – | – | – |
| ACI(MF) | 21,400 | 19.5096 | 15,152 | 0.2320 | Extended dominated | MF(ACI) |
| ACI(ACI) | 22,461 | 19.5363 | 1062 | 0.0267 | 62,658 | MF(ACI) |
| Probabilistic: discounted | | | | | | |
| MF(ACI) | 6250 | 19.2966 | – | – | – | – |
| ACI(MF) | 21,438 | 19.5140 | 15,188 | 0.2290 | Extended dominated | MF(ACI) |
| ACI(ACI) | 22,512 | 19.5484 | 1074 | 0.0384 | 64,581 | MF(ACI) |

**FIGURE 41** Cost-effectiveness acceptability curve: utility = 0.817.

Utility for choose no second repair set to mid-point of success and failure

- Data used for ACI failure rates: Nawaz *et al.*⁸⁰ (whole cohort).
- Data used for MF failure rates: pooled data (Layton 2015,¹¹⁴ Knutsen 2007,⁶⁷ Saris 2009⁷⁰).
- Cost of harvesting: £870.
- Cost of implantation: £2396.
- Cost of cells: £16,000.

Utility for those who choose no second repair: 0.746. This also reduces the marginal QALY gains from ACI as first procedure because the larger proportion who do not do well after MF have their utility increased.

Table 42 shows the results when the utility value for those who choose no second repair is 0.746. The deterministic discounted results highlighted that the option of ACI(MF) was extendedly dominated and therefore the ICER comparing ACI(ACI) with MF(ACI) was just over £27,000. Figure 42 shows that at a willingness to pay of £26,000 there is not much difference between the two remaining options.

TABLE 42 Deterministic and probabilistic cost-effectiveness results: utility = 0.746

| Procedure | Total mean costs (£) | Total mean QALYs | Incremental costs (£) | Incremental QALYs | ICER vs. MF(ACI) | Comparator |
|------------------------------------|----------------------|------------------|-----------------------|-------------------|--------------------|------------|
| Deterministic: undiscounted | | | | | | |
| MF(ACI) | 7736 | 36.4022 | – | – | – | – |
| ACI(MF) | 22,661 | 37.2311 | 14,926 | 0.8289 | Extended dominated | MF(ACI) |
| ACI(ACI) | 24,134 | 37.3317 | 1473 | 0.1006 | 17,643 | MF(ACI) |
| Deterministic: discounted | | | | | | |
| MF(ACI) | 6248 | 18.0702 | – | – | – | – |
| ACI(MF) | 21,400 | 18.6197 | 15,152 | 0.5495 | Extended dominated | MF(ACI) |
| ACI(ACI) | 22,461 | 18.6680 | 1062 | 0.0483 | 27,123 | MF(ACI) |
| Probabilistic: discounted | | | | | | |
| MF(ACI) | 6248 | 18.0400 | – | – | – | – |
| ACI(MF) | 21,419 | 18.6257 | 15,171 | 0.5857 | Extended dominated | MF(ACI) |
| ACI(ACI) | 22,496 | 18.6684 | 1077 | 0.0427 | 25,857 | MF(ACI) |

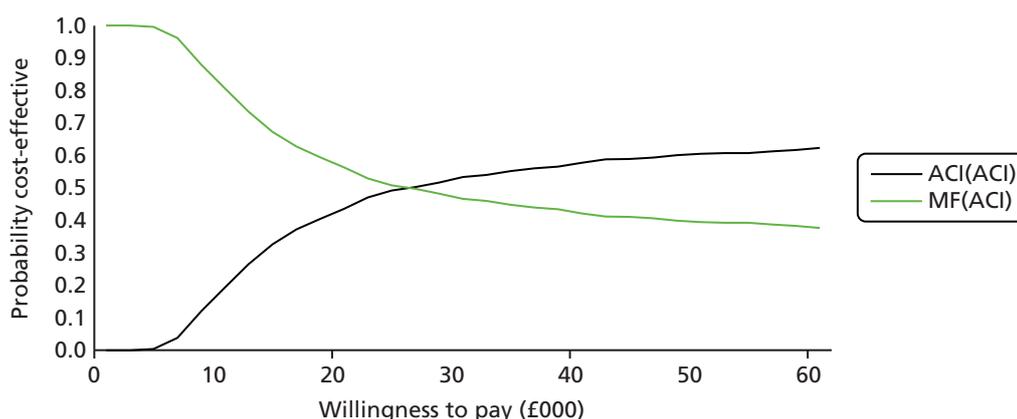


FIGURE 42 Cost-effectiveness acceptability curve: utility = 0.746.

Subgroup analyses

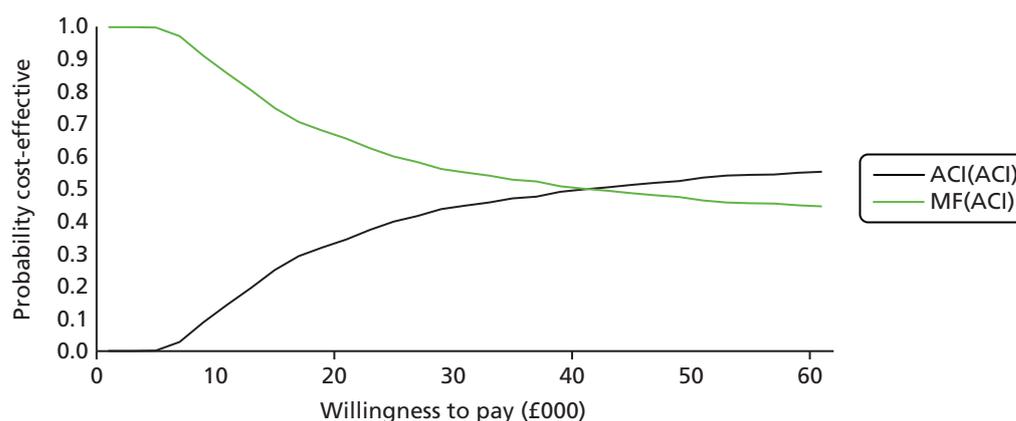
Individuals with prior repair attempts

- Data used for ACI failure rates: Nawaz *et al.*⁸⁰ previous intervention.
- Data used for MF failure rates: pooled data (Layton 2015,¹¹⁴ Knutsen 2007,⁶⁷ Saris 2009⁷⁰).
- Cost of harvesting: £870.
- Cost of implantation: £2396.
- Cost of cells: £16,000.

Table 43 shows the results for those individuals with prior repair attempts. The deterministic discounted results highlighted that the option of ACI(MF) was extendedly dominated and therefore the ICER comparing ACI(ACI) with MF(ACI) was just over £38,000. Figure 43 shows that, at a willingness to pay of £30,000 or less, MF(ACI) is the more cost-effective option.

TABLE 43 Deterministic and probabilistic cost-effectiveness results: previous intervention

| Procedure | Total mean costs (£) | Total mean QALYs | Incremental costs (£) | Incremental QALYs | ICER vs. MF(ACI) | Comparator |
|------------------------------------|----------------------|------------------|-----------------------|-------------------|--------------------|------------|
| Deterministic: undiscounted | | | | | | |
| MF(ACI) | 7736 | 34.2885 | – | – | – | – |
| ACI(MF) | 22,718 | 34.7835 | 14,983 | 0.4950 | Extended dominated | MF(ACI) |
| ACI(ACI) | 24,314 | 34.9315 | 1595 | 0.1480 | 25,780 | MF(ACI) |
| Deterministic: discounted | | | | | | |
| MF(ACI) | 6248 | 17.1350 | – | – | – | – |
| ACI(MF) | 21,462 | 17.4918 | 15,214 | 0.3569 | Extended dominated | MF(ACI) |
| ACI(ACI) | 22,746 | 17.5661 | 1284 | 0.0743 | 38,262 | MF(ACI) |
| Probabilistic: discounted | | | | | | |
| MF(ACI) | 6236 | 17.1315 | – | – | – | – |
| ACI(MF) | 21,503 | 17.4889 | 15,267 | 0.3575 | Extended dominated | MF(ACI) |
| ACI(ACI) | 22,798 | 17.5522 | 1295 | 0.0632 | 39,370 | MF(ACI) |

**FIGURE 43** Cost-effectiveness acceptability curve (previous interventions).

Individuals without prior repair attempts

- Data used for ACI failure rates: Nawaz *et al.*⁸⁰ no previous intervention.
- Data used for MF failure rates: pooled data (Layton 2015,¹¹⁴ Knutsen 2007,⁶⁷ Saris 2009⁷⁰)
- Cost of harvesting: £870.
- Cost of implantation: £2396.
- Cost of cells: £16,000.

Table 44 shows the results for those individuals without any prior repair attempts. The deterministic discounted results showed that the ICER comparing ACI(ACI) with ACI(MF) was just over £15,500. Figure 44 shows that at a willingness to pay of £12,000 there is not much difference between the two remaining options.

TABLE 44 Deterministic and probabilistic cost-effectiveness results: no previous intervention

| Procedure | Total mean costs (£) | Total mean QALYs | Incremental costs (£) | Incremental QALYs | ICER vs. MF(ACI) | Comparator |
|------------------------------------|----------------------|------------------|-----------------------|-------------------|------------------|------------|
| Deterministic: undiscounted | | | | | | |
| MF(ACI) | 7736 | 34.2885 | – | – | – | – |
| ACI(MF) | 21,956 | 37.4216 | 14,220 | 3.1332 | 4539 | MF(ACI) |
| ACI(ACI) | 22,826 | 37.5038 | 870 | 0.0822 | 10,586 | ACI(MF) |
| Deterministic: discounted | | | | | | |
| MF(ACI) | 6248 | 17.1350 | – | – | – | – |
| ACI(MF) | 21,101 | 18.7446 | 14,853 | 1.6097 | 9227 | MF(ACI) |
| ACI(ACI) | 21,644 | 18.7793 | 543 | 0.0347 | 15,659 | ACI(MF) |
| Probabilistic: discounted | | | | | | |
| MF(ACI) | 6268 | 17.1506 | – | – | – | – |
| ACI(MF) | 21,114 | 18.6100 | 14,846 | 1.4594 | 10,172 | MF(ACI) |
| ACI(ACI) | 21,930 | 18.6211 | 816 | 0.0310 | 26,324 | ACI(MF) |

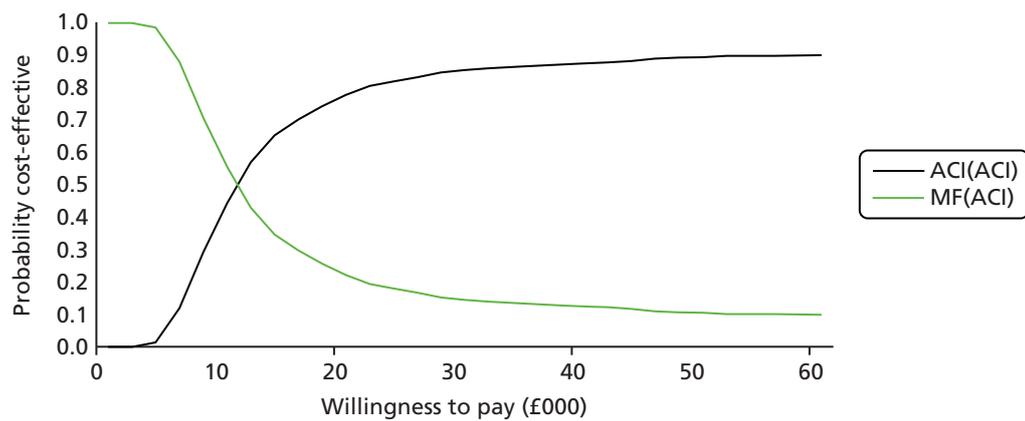


FIGURE 44 Cost-effectiveness acceptability curve: no previous interventions.

Individuals with Kellgren–Lawrence grade 0

- Data used for ACI failure rates: Nawaz *et al.*⁸⁰ Kellgren–Lawrence grade 0.
- Data used for MF failure rates: pooled data (Layton 2015,¹¹⁴ Knutsen 2007,⁶⁷ Saris 2009⁷⁰).
- Cost of harvesting: £870.
- Cost of implantation: £2396.
- Cost of cells: £16,000.

Table 45 shows the results for those individuals with Kellgren–Lawrence grade 0. The deterministic discounted results highlighted that the ICER comparing ACI(ACI) with ACI(MF) was just over £15,500.

Individuals with Kellgren–Lawrence grade 1

- Data used for ACI failure rates: Nawaz *et al.*⁸⁰ Kellgren–Lawrence grade 1.
- Data used for MF failure rates: pooled data (Layton 2015,¹¹⁴ Knutsen 2007,⁶⁷ Saris 2009⁷⁰).
- Cost of harvesting: £870.
- Cost of implantation: £2396.
- Cost of cells: £16,000.

Table 46 shows the results for those individuals with Kellgren–Lawrence grade 1. The deterministic discounted results highlighted that the option of ACI(MF) was extendedly dominated and therefore the ICER comparing ACI(ACI) with MF(ACI) was just over £17,000.

Individuals with Kellgren–Lawrence grade 2

- Data used for ACI failure rates: Nawaz *et al.*⁸⁰ Kellgren–Lawrence grade 2.
- Data used for MF failure rates: pooled data (Layton 2015,¹¹⁴ Knutsen 2007,⁶⁷ Saris 2009⁷⁰).
- Cost of harvesting: £870.
- Cost of implantation: £2396.
- Cost of cells: £16,000.

Table 47 shows the results for those individuals with Kellgren–Lawrence grade 2. The deterministic discounted results highlighted that the option of ACI(MF) was extendedly dominated and therefore the ICER comparing ACI(ACI) with MF(ACI) was just over £20,000.

TABLE 45 Deterministic cost-effectiveness results: Kellgren–Lawrence grade 0

| Procedure | Total mean costs (£) | Total mean QALYs | Incremental costs (£) | Incremental QALYs | ICER vs. MF(ACI) | Comparator |
|------------------------------------|----------------------|------------------|-----------------------|-------------------|------------------|------------|
| Deterministic: undiscounted | | | | | | |
| MF(ACI) | 7736 | 34.2885 | – | – | – | – |
| ACI(MF) | 22,489 | 36.4611 | 14,753 | 2.1726 | 6791 | MF(ACI) |
| ACI(ACI) | 23,727 | 36.5794 | 1238 | 0.1183 | 10,470 | ACI(MF) |
| Deterministic: discounted | | | | | | |
| MF(ACI) | 6248 | 17.1350 | – | – | – | – |
| ACI(MF) | 21,294 | 18.3745 | 15,046 | 1.2395 | 12,138 | MF(ACI) |
| ACI(ACI) | 22,079 | 18.4247 | 785 | 0.0503 | 15,618 | ACI(MF) |

TABLE 46 Deterministic cost-effectiveness results: Kellgren–Lawrence grade 1

| Procedure | Total mean costs (£) | Total mean QALYs | Incremental costs (£) | Incremental QALYs | ICER vs. MF(ACI) | Comparator |
|------------------------------------|----------------------|------------------|-----------------------|-------------------|--------------------|------------|
| Deterministic: undiscounted | | | | | | |
| MF(ACI) | 7736 | 34.2885 | – | – | – | – |
| ACI(MF) | 22,679 | 35.7135 | 14,943 | 1.4250 | 10,486 | MF(ACI) |
| ACI(ACI) | 24,129 | 35.8516 | 1450 | 0.1381 | 10,499 | ACI(MF) |
| Deterministic: discounted | | | | | | |
| MF(ACI) | 6248 | 17.1350 | – | – | – | – |
| ACI(MF) | 21,395 | 18.0173 | 15,147 | 0.8824 | Extended dominated | MF(ACI) |
| ACI(ACI) | 22,408 | 18.0798 | 1013 | 0.0624 | 17,104 | MF(ACI) |

TABLE 47 Deterministic cost-effectiveness results: Kellgren–Lawrence grade 2

| Procedure | Total mean costs (£) | Total mean QALYs | Incremental costs (£) | Incremental QALYs | ICER vs. MF(ACI) | Comparator |
|------------------------------------|----------------------|------------------|-----------------------|-------------------|--------------------|------------|
| Deterministic: undiscounted | | | | | | |
| MF(ACI) | 7736 | 34.2885 | – | – | – | – |
| ACI(MF) | 22,718 | 35.4402 | 14,983 | 1.1517 | Extended dominated | MF(ACI) |
| ACI(ACI) | 24,233 | 35.5842 | 1514 | 0.1440 | 12,732 | MF(ACI) |
| Deterministic: discounted | | | | | | |
| MF(ACI) | 6248 | 17.1350 | – | – | – | – |
| ACI(MF) | 21,423 | 17.8779 | 15,175 | 0.7430 | Extended dominated | MF(ACI) |
| ACI(ACI) | 22,520 | 17.9447 | 1097 | 0.0667 | 20,096 | MF(ACI) |

Individuals with Kellgren–Lawrence grade 3

- Data used for ACI failure rates: Nawaz *et al.*⁸⁰ Kellgren–Lawrence grade 3.
- Data used for MF failure rates: pooled data (Layton 2015,¹¹⁴ Knutsen 2007,⁶⁷ Saris 2009⁷⁰).
- Cost of harvesting: £870.
- Cost of implantation: £2396.
- Cost of cells: £16,000.

Table 48 shows the results for those individuals with Kellgren–Lawrence grade 3. The deterministic discounted results highlighted that the option of ACI(MF) was extendedly dominated and therefore the ICER comparing ACI(ACI) with MF(ACI) was just over £21,000.

Pooled autologous chondrocyte implantation curve (six studies)

- Data used for ACI failure rates: pooled data (Knutsen 2007,⁶⁷ Minas 2014,¹³⁶ Moseley 2010,¹¹⁵ Nawaz,¹⁶⁵ Niemeyer 2014,¹¹⁷ Vanlauwe 2011⁴³).
- Data used for MF failure rates: pooled data (Layton 2015,¹¹⁴ Knutsen 2007,⁶⁷ Saris 2009⁷⁰).
- Cost of harvesting: £870.
- Cost of implantation: £2396.
- Cost of cells: £16,000.

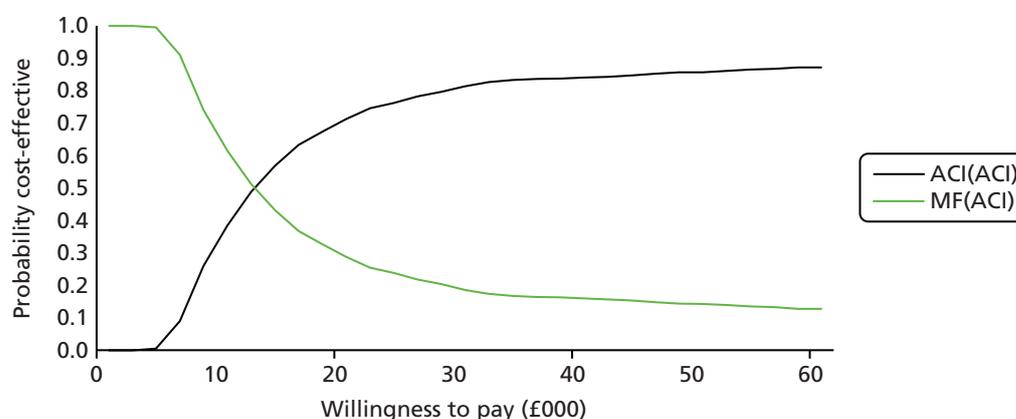
TABLE 48 Deterministic cost-effectiveness results: Kellgren–Lawrence grade 3

| Procedure | Total mean costs (£) | Total mean QALYs | Incremental costs (£) | Incremental QALYs | ICER vs. MF(ACI) | Comparator |
|------------------------------------|----------------------|------------------|-----------------------|-------------------|--------------------|------------|
| Deterministic: undiscounted | | | | | | |
| MF(ACI) | 7736 | 34.2885 | – | – | – | – |
| ACI(MF) | 22,726 | 35.3609 | 14,990 | 1.0724 | Extended dominated | MF(ACI) |
| ACI(ACI) | 24,258 | 35.5063 | 1532 | 0.1455 | 13,566 | MF(ACI) |
| Deterministic: discounted | | | | | | |
| MF(ACI) | 6248 | 17.1350 | – | – | – | – |
| ACI(MF) | 21,430 | 17.8358 | 15,183 | 0.7008 | Extended dominated | MF(ACI) |
| ACI(ACI) | 22,552 | 17.9038 | 1122 | 0.0680 | 21,207 | MF(ACI) |

Table 49 shows the results for the pooled six ACI data sets. The deterministic discounted results showed that the ICER comparing ACI(ACI) with MF(ACI) was just under £12,000. Figure 45 shows that, at a willingness to pay of £12,000 or less, MF(ACI) is the more cost-effective option.

TABLE 49 Deterministic and probabilistic cost-effectiveness results: six ACI data sets

| Procedure | Total mean costs (£) | Total mean QALYs | Incremental costs (£) | Incremental QALYs | ICER vs. MF(ACI) | Comparator |
|------------------------------------|----------------------|------------------|-----------------------|-------------------|------------------|------------|
| Deterministic: undiscounted | | | | | | |
| MF(ACI) | 7736 | 34.2885 | – | – | – | – |
| ACI(MF) | 22,140 | 36.7771 | 14,405 | 2.4886 | 5788 | MF(ACI) |
| ACI(ACI) | 23,195 | 37.8748 | 1055 | 0.0978 | 10,794 | ACI(MF) |
| Deterministic: discounted | | | | | | |
| MF(ACI) | 6248 | 17.1350 | – | – | – | – |
| ACI(MF) | 21,192 | 18.4290 | 14,944 | 1.2940 | 11,549 | MF(ACI) |
| ACI(ACI) | 21,933 | 18.4734 | 741 | 0.0444 | 16,708 | ACI(MF) |
| Probabilistic: discounted | | | | | | |
| MF(ACI) | 6271 | 17.1731 | – | – | – | – |
| ACI(MF) | 21,235 | 18.4253 | 14,964 | 1.2522 | 11,950 | MF(ACI) |
| ACI(ACI) | 21,991 | 18.4948 | 757 | 0.0695 | 10,882 | ACI(MF) |

**FIGURE 45** Cost-effectiveness acceptability curve: six ACI data sets.

Using utility data from Vericel

- Data used for ACI failure rates: Nawaz *et al.*⁸⁰ (whole cohort)
- Data used for MF failure rates: pooled data (Layton 2015,¹¹⁴ Knutsen 2007,⁶⁷ Saris 2009⁷⁰)

Tables 50 and 51 shows the results when using utility data for Vericel. The deterministic discounted results highlighted that the option of ACI(MF) was extendedly dominated and therefore the ICER comparing ACI(ACI) with MF(ACI) was just over £15,500. Figure 46 shows that at a willingness to pay of £16,000 there is not much difference between the two remaining options.

TABLE 50 Utility data from Vericel

| Time point | MACI | MF |
|-----------------------------|-----------------|-----------------|
| Baseline | | |
| <i>n</i> | 141 | |
| Mean utility value (SD) | 0.484 (0.296) | |
| Response at week 52 | | |
| <i>n</i> | 71 | 68 |
| Mean utility value (SD) | 0.7848 (0.2113) | 0.7472 (0.2270) |
| Response at week 104 | | |
| <i>n</i> | 70 | 70 |
| Mean utility value (SD) | 0.8051 (0.1899) | 0.7188 (0.2969) |
| Response at week 156 | | |
| <i>n</i> | 65 | 59 |
| Mean utility value (SD) | 0.8131 (0.2105) | 0.7769 (0.2553) |

TABLE 51 Deterministic and probabilistic cost-effectiveness results

| Procedure | Total mean costs (£) | Total mean QALYs | Incremental costs (£) | Incremental QALYs | ICER vs. MF (ACI) | Comparator |
|------------------------------------|----------------------|------------------|-----------------------|-------------------|--------------------|------------|
| Deterministic: undiscounted | | | | | | |
| MF(ACI) | 7736 | 33.8297 | – | – | – | – |
| ACI(MF) | 22,661 | 35.2364 | 14,926 | 1.4067 | Extended dominated | MF(ACI) |
| ACI(ACI) | 24,134 | 35.3784 | 1473 | 0.1420 | 10,588 | MF(ACI) |
| Deterministic: discounted | | | | | | |
| MF(ACI) | 6248 | 16.6956 | – | – | – | – |
| ACI(MF) | 21,400 | 17.6627 | 15,152 | 0.9671 | Extended dominated | MF(ACI) |
| ACI(ACI) | 22,461 | 17.7317 | 1061 | 0.0690 | 15,648 | MF(ACI) |
| Probabilistic: discounted | | | | | | |
| MF(ACI) | 6283 | 16.7221 | – | – | – | – |
| ACI(MF) | 21,381 | 17.6499 | 15,098 | 0.9277 | Extended dominated | MF(ACI) |
| ACI(ACI) | 22,456 | 17.7528 | 1075 | 0.1029 | 15,692 | MF(ACI) |

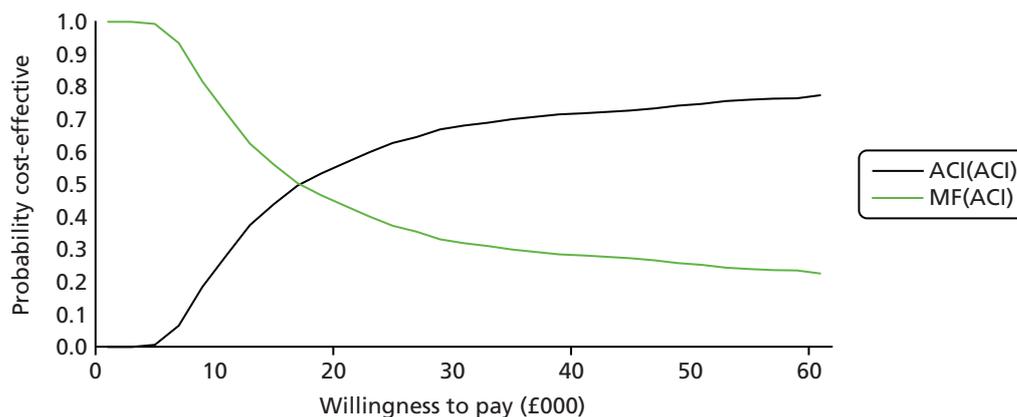


FIGURE 46 Cost-effectiveness acceptability curve.

Summary of long-study modelling

- In many scenarios, ACI(MF) is extended dominated, meaning that the relevant choice is between ACI(ACI) and MF(ACI), and the use of MF as a post-ACI treatment is not a relevant alternative.
- The exceptions to this tend to be in scenarios in which ACI is particularly effective (e.g. no previous repair, Kellgren–Lawrence grade of 0), for which there is less additional benefit to be gained from a more effective second procedure.
- Decreases in ACI treatment costs, unsurprisingly, lead to reductions in the ICERs for ACI.
- Higher utilities in the 'no further treatment' state make ACI less cost-effective, as there is less benefit gained from successful procedures, and, likewise, lower utilities in the 'no further treatment' state make ACI more cost-effective.
- Including evidence from a wider range of studies makes ACI appear more cost-effective than using data from Nawaz *et al.*⁸⁰ alone.

Chapter 7 Discussion

Statement of principal findings:

- ACI has evolved since the last review by NICE in 2005, and now chondrocytes are seeded into membranes or scaffolds, rather than a liquid suspension of cells being secured under a periosteal or collagen cap.
- Selection of the chondrocytes that are most likely to produce good quality repairs ('characterisation') is now used, though there are no published trials proving benefit.
- ACI is an effective way of treating defects in articular cartilage, giving good results in over 80% of patients. If results are good at 2 years, benefit is generally sustained for up to 10 years. A very large UK cohort showed graft survivals of 78% at 5 years, and 51% at 10 years.
- The main comparator, MF, is effective in a smaller proportion and appears to be less durable.
- Our economic modelling found that ACI appeared to be cost-effective compared with MF, with a key driver being duration of benefit and likely avoidance or postponement of a second repair or of knee replacement. MF was less costly, but provided fewer QALYs.
- Total costs were influenced by the proportion needing a second repair, and by the method used for second repairs. If all second repairs were by ACI, the cost per QALY gained for initial ACI compared with initial MF was £8925. If all second repairs were by MF, the cost per QALY gained was £9788. These results were confirmed by the CEACs: so if the decision-maker is willing to pay £20,000 for a QALY, ACI is 56–59% more likely to be cost-effective than MF. For both scenarios, ACI as a first repair was more cost-effective than MF as a first repair.
- There is a shortage of long-term studies, particularly of MF. As requested by NICE, we carried out survival analysis making the best of what data there were. Caveats are required.
- We included six studies of long-term results of ACI, the best of which was by Nawaz *et al.*,⁸⁰ from Stanmore. It was best because of size (827 – all of the other studies put together provided 371); it reflected UK practice (albeit from a centre of excellence); it provided data from the period 1998–2008, on different generations of ASCI; and it provided very useful subgroup data.
- Using the older data, MF comes out less well, with progressive failure over time.
- As noted in the previous report, ACI is less successful in people who had had prior repair attempts, such as MF.
- The new base-case analysis used MF followed, if necessary, by ACI, as the lowest cost option, with other options being compared with that. The ICER for ACI as a primary procedure compared with MF was around £19,000. The range of economics analyses produced ICERs that might be considered acceptable by NICE.

Strengths and weaknesses of evidence

- At the last appraisal, there were no long-term data from trials. The evidence base has also evolved with data on longer-term follow-up both from trials and cohort studies. However, the longest-term data come from older generations of ACI, and recruits to such studies had often had several prior attempts at repair, which appear to reduce the effectiveness of ACI.
- Because of short follow-up of the MACI trials, there is a lack of long-term utility data.
- The TIG/ACT trial⁴³ of ChondroCelect used ACI-P, which has now been superseded by ACI-C or MACI. ChondroCelect cells are now used in a MACI procedure wherein the cells are loaded on to a membrane by the surgeon.
- There is a general problem when long-term results are needed but the technology continues to evolve. Data on long-term results come mainly from first-generation ACI.
- Utilities vary considerably among studies. For example, baseline utility before repair ranges from 0.41 (Derrett2005¹⁴⁰) to 0.532 or 0.504 (ACTIVE,³⁵ MF and ACI groups, respectively) to 0.654 (Gerlier 2010¹⁴¹).

- When considering survival curves extrapolated beyond the observed data, it should be borne in mind that the extrapolation assumes that the curve based on the observed data will continue. However, this may not always be the case. For example, if ACI failures occurred mainly in the early observed years, longer-term observations would show a levelling off. However, this may apply only after successful ACI. Bhosale *et al.*,¹⁰⁸ from Oswestry, in a series of 80 patients, reported that success at 15 months was sustained, but average follow-up was for only 5 years. The Nawaz study⁸⁰ suggests that when ACI is most successful, the survival curve shows some levelling off by about 7 years, whereas in those in whom it fails, the curve shows a linear decline.
- The lack of data on the benefits of MF compared with debridement alone is a problem. (And it is worth remembering that in a previous assessment report, we noted a lack of evidence for debridement and lavage over non-operative approaches.³)
- We relied heavily on the Nawaz study.⁸⁰ We confirmed with the lead author that the patients in the ACI arm of the Bentley trial,⁷⁸ and the cohort in the long-term outcome study by Biant *et al.*,⁷⁹ were included. Before obtaining that information, we had included the Bentley⁷⁸ and Biant⁷⁹ studies on pooled survival analysis. Curiously, removing them worsened the ACI results, despite them having, in some ways, patients with poorer prognostic factors. For example, the proportions having previous repair attempts were 34% in Nawaz, 94% in Bentley and 73% in Biant. The patients in the Bentley⁷⁸ and Biant⁷⁹ studies were from the earliest days (1998–2001) and were ‘salvage’ cases after means of 1.7 and 1.3 previous procedures, respectively.
- The reason for the better results in the Minas series¹³⁶ than in the Nawaz study⁸⁰ is not clear. The Minas¹³⁶ patients all had MACI. The definitions of failure may explain some of the difference, with failure in Minas¹³⁶ very surgically defined, such that some failures in the Nawaz study might not have been classed as failure by Minas *et al.*¹³⁶
- Another variable that may cause differences in outcomes could be differences in comparator treatments, such as drilling and MF. After MF, microscopic cracks form around the holes. These do not occur when bone is drilled. So MF may do more damage to the subchondral bone.
- As noted, there are rather more long-term studies of ACI than of MF. Why are there so few of MF? Could it be that long-term results are poor and that people with data do not publish it? Should the questions in this appraisal have included: Should MF be done at all, irrespective of whether ACI is available?
- The evidence base has many deficiencies. One is that older studies tended to recruit patients who had had previous attempts at repair, and these may give a misleadingly pessimistic picture of how ACI would perform if used as first procedure.

Other issues

Asymptomatic lesions

The BASK UK³⁹ Consensus recommends that patients should have conservative treatment with physical therapy before being considered for cartilage repair. Many will become asymptomatic and will no longer qualify for ACI according to the NICE scope. However, their cartilage defect will not recover spontaneously, and they are likely to develop OA in later years. Should they be considered for ACI?

The Dutch Orthopaedic Association recommends treatment of asymptomatic ICRS grade 5 lesions.²²

Osteoarthritis

The NICE scope excludes people with ‘advanced osteoarthritis’. OA can be defined as generalised degenerative change affecting both sides of an articulation. ACI is used for isolated cartilage defects. There can be isolated defects on both surfaces (‘kissing lesions’), which could be considered for ACI if the rest of the joint is in good order. Patients with only early OA (less than grade 2 which has definite osteophytes and possible joint space narrowing) could have been included in some trials. However, no details for such a subgroup are given in the results. In the TIG/ACT trial,⁴³ patients with advanced OA (as defined by radiographic atlas OA grade 2–3) were excluded.

A systematic review of cartilage repair in early OA by de Windt *et al.*¹⁶⁶ found evidence of benefit in those having various forms of ACI, ranging from ACI-P to MACI. Early OA was defined in different ways in the nine case series, and de Windt *et al.*¹⁶⁶ described the studies as being of 'generally low methodological quality'. Nevertheless, they reported that outcomes to 9 years were good, and suggested that ACI in early OA might be used to postpone TKR, but recommended a RCT.

There may, therefore, be a place for ACI in early OA, even if only to postpone TKR until patients are older, and some of the ICERs reported earlier are within the acceptable range. However, the evidence base is much weaker than for purely chondral lesions.

However, if ACI were to be restricted based on radiological signs of OA, there are some problems to be considered. One of the difficulties in comparing the results of studies involving patients with OA is the definition of the disease and the assessment of its severity. The European League Against Rheumatism (EULAR) definition of OA emphasises the importance of pain and functional loss alongside physical changes in the joint, but this definition is hard to objectively apply in research terms, and symptoms are significantly influenced by environmental and psychosocial factors.^{167–170}

There is a variable relationship between symptoms and structural changes in OA and it is recognised that plain radiographs, MRI and arthroscopic findings do not universally correlate with pain or physical function.^{171–173}

The most common method for assessing structural changes in knee OA is plain radiography, graded using the Kellgren–Lawrence classification.¹⁷⁴ Care has to be taken in interpreting plain radiographic findings, as Kellgren–Lawrence grades have moderate, but not strong, correlations with other measures of structural change, such as MRI measures of OA or operative findings.^{175–180}

The Kellgren–Lawrence classification is a widely accepted tool in OA research and good reliability has been quoted in series in which the assessors were experienced in its use.^{173,178} However, it is based on a subjective assessment of structural changes and different authors often apply different criteria to define the boundaries between the grades, making comparisons across studies difficult.¹⁸¹

The boundary between Kellgren–Lawrence grades 2 and 3 is often difficult to define, as the interpretation of 'possible' and 'definite' joint space narrowing can be very subjective.¹⁸² However, this is not so important when considering suitability for ACI, as the Nawaz study⁸⁰ showed that there was little difference in outcomes. The distinction between lower Kellgren–Lawrence grades is also difficult and is dependent on the interpretation of small osteophytes, which can variably give a score of 0, 1 or 2, depending on the exact definitions used and the radiological technique.¹⁸¹

The diagnosis of OA is often made based on the combination of symptoms and a Kellgren–Lawrence grade of '2 or more', despite evidence that Kellgren–Lawrence grade of 1 ('doubtful osteophytes') has a high chance of progressing to '2 or more' with time.^{183,184}

The studies in this review have varied in terms of their reporting of the radiological assessment, and definitions were not always clearly defined in the reports, which may explain some of the variance in findings between studies. For example, relatively little detail is given in the Minas paper¹⁸⁵ on the radiological assessment and the Kellgren–Lawrence paper is not referenced, whereas the radiological grading is reported in detail by Nawaz *et al.*⁸⁰ A relatively high proportion of cases with Kellgren–Lawrence grade 2 or above were reported by Knutsen,⁶⁷ which may explain the poor results for ACI in this series in comparison to others.

Defining OA is problematic. A big cartilage lesion with pain and some joint space loss could variably be defined as no, mild or moderate OA.

Age threshold for knee replacement

In our modelling we have assumed that TKR would not be performed for people with OA until age 55 years or later. We used that age restriction because knee replacements do not last for ever, and replacing a replacement is more difficult, more expensive and less successful than the first replacement, and may not last as long.

With increasing longevity, it may no longer be the case that a knee replacement in someone over 60 years is likely to last them all of their days. Perhaps especially in women who live longer. However, a TKR in a younger person with OA is very likely to need replacement. (This may not apply to people having knee replacement because of inflammatory arthritis because their activity, and hence the stresses put upon the prosthesis, will often be limited by problems with other joints.)

In the National Joint Registry 2015 report,¹⁵⁷ figure 3.16 shows that the probability of a first revision after TKR is higher in people who have replacements at younger ages. Those who have TKR under the age of 55 years have a 12% probability of it being replaced by 11 years, which is more than double the risk after first TKR at older ages.

It is therefore a major decision to carry out TKR in people with OA under the age of 60 years, and very few are done. It should be noted that TKR is rarely an absolute necessity. The aim is to reduce pain, and that can be done in other ways, such as with analgesics or reducing activity.

It should also be borne in mind that TKR does not fully restore knee function. The TKR does not move like a normal knee, and younger, active patients may find function on stairs and slopes disappointing.

Autologous chondrocyte implantation can restore normal function in younger patients. In patients who are older but too young for TKR, but who do not have generalised wear and tear, ACI may help bridge the gap to TKR, even if the results are not as good as in younger patients with only an isolated chondral defect.

Body mass index

Jaiswal *et al.*¹⁸⁶ from Stanmore reported a lack of benefit from ACI or MACI in patients with BMI score of over 30 kg/m², though this was based on small numbers in the high BMI group. Their data came from the trial of MACI versus ACI. In 53 patients with BMI scores of under 25 kg/m², 82% of patients had a good or excellent result. In the overweight group (BMI 25–30 kg/m²) 49% (22 of 45) had a good or excellent result, whereas only one of 18 patients with BMI over 30 kg/m² had a good result.

Mithöfer *et al.*⁸² also reported worse outcomes in those with BMI score of over 30 kg/m². Behery *et al.*⁸³ reported no correlation but had data on only eight patients.

Data on the effect of high BMI on outcomes of cartilage repair is sparse. Jaiswal *et al.*¹⁸⁶ reported that their literature review found few previous studies. In most studies, mean BMI scores were well below 30 kg/m², perhaps because cartilage injuries occur largely in people active in sports. Jaiswal used the term 'obese' but some sportsmen with high BMI scores may be lean but very muscular.

Similar findings have been reported for MF by Asik *et al.*,¹⁰⁶ with better results in those with BMI score of less than 25 kg/m².

Research needs

Recommendations for research made in the systematic reviews.

Some of the recommendations made in the reviews are now out of date and are not included here. Other recommendations include:

- High-quality clinical trials are needed, fulfilling the following criteria:
 - Multicentre, adequate sample size with long-term follow-up (preferably 5–10 years).
 - Patients in trials should be stratified based on BMI, defect location, post-debridement defect size and previous cartilage repair.
 - Transparent patient enrolment with clearly stated inclusion and exclusion criteria.
 - Proper independently performed randomisation techniques.
 - No concurrent surgical interventions (anterior cruciate ligament reconstruction, realignment osteotomy, meniscal surgery, etc.), consistent surgical technique.
 - Use of validated, responsive and reliable patient-orientated outcome measures; clear reporting of data with a statement of both clinical relevance and significance; and use of independent assessors.
 - Further information is needed on the relationship between clinical, histological and radiological outcomes, and the most appropriate measure of functional outcomes that relates to a generic measure of health-related QoL.
- Cohort studies of long-term effects (≥ 10 years) are needed.
- Research is needed to explain lack of return to sports by some patients.
- Prospective long-term studies are needed to determine if articular cartilage repair in athletes can influence the high incidence of OA associated with high-impact sports.
- More studies should be done on the maturation process of finally formed repair tissue and on appropriate rehabilitation programmes for the different techniques.

Fourth-generation autologous chondrocyte implantation

There are several lines of investigation.

Mesenchymal cells

It has been suggested that mesenchymal stem cells from bone marrow can be used as an alternative to ACI and that their reproduction is less affected by age. (For reviews see Nakamura *et al.*¹⁸⁷ and Perera *et al.*¹⁸⁸)

A review of scaffold-based repair by Filardo *et al.*¹⁸⁹ mentions another option, using mesenchymal stem cells and a degradable scaffold, covered with fibrin.¹⁹⁰

The ASCOT trial¹⁹¹ will compare repairs with chondrocytes and bone marrow mesenchymal stem cells, and with the combination of both.

INSTRUCT

This appears to be a one-stage procedure mixing chondrocytes and bone marrow cells, without cell culture. Cells from a biopsy of cartilage are mixed with bone marrow cells then seeded into a porous scaffold, which is then implanted into the defect. Evidence comes from a poster by Hendriks *et al.*¹⁹² So far, only 37 patients had reached 12-month follow-up, of whom 72% had hyaline cartilage on biopsy.

Cartilage implantation

The development here is that instead of implanting cultured chondrocytes into the defect, the autologous chondrocytes are used to grow new cartilage in the laboratory, which is then implanted.¹⁹³

Gel-type autologous chondrocyte implantation

Gel-type ACI appears to be a new variant without using membrane or periosteum, but using cells held in place with fibrin. Choi *et al.*¹⁹⁴ report a case series with 98 patients. There do not appear to have been RCTs against standard ACI.

Single-stage procedures

Cole *et al.*⁸⁹ report a RCT with 29 patients, comparing MF (nine patients) with a cartilage autograft implantation system (CAIS) in which chondrocytes are not sent for culture. Instead, hyaline cartilage is harvested in similar amount as for traditional ACI, but then minced and attached to a biodegradable scaffold with fibrin glue, in a single operation. Results at 24 months showed some advantages for the CAIS group, with IKDC score 83 for CAIS and 60 for MF, and KOOS also better.

Other cells

Mizuno *et al.*¹⁹⁵ report that ear cartilage cells can be used, at least in dogs.

New forms of microfracture

Filardo *et al.*¹⁸⁹ report five case series of autologous matrix-induced chondrogenesis, which combines MF with a collagen matrix to stabilise the blood clot. Long-term results are not yet available.

Siclari *et al.*¹⁹⁶ used a combination of MF and a cell-free hyaluronan cap that had been immersed in autologous plasma in 52 patients. At 2 years, KOOS results showed good improvement. Biopsies were taken from four patients and showed hyaline or hyaline-like repair tissue.

Metal or plastic patches for knees

These were excluded by NICE as comparators, but sound sufficiently promising to be used in trials. They may not be suitable for younger patients but might be an option for the 40–60 years subgroup, perhaps in order to postpone knee replacement.

The HemiCAP® (Arthrosurface, Franklin, MA, USA)¹⁹⁷ is used for resurfacing localised damage in femoral condyles, and is described by the manufacturer as a ‘contoured articular resurfacing implant’, and as ‘bridging the gap between biological therapies and TKR’. The evidence base seems to consist of a few case series with no RCTs:

- Patello Femoral HemiCAP®¹⁹⁸
- UniCAP®¹⁹⁹
- HemiCAP® Classic.²⁰⁰

The BioPoly™ RS Knee System²⁰¹ (Schwartz Biomedical, Fort Wayne, IN, USA) is CE marked for sale in the European Union. It is a hyaluronic and polythene implant for repairing the joint surface.

The Episealer (Episurf AB, Stockholm, Sweden)²⁰² comes in two forms, for femoral condyle and trochlea, and is described as a small metallic button with implants tailored for each patient.

These products are said to allow rapid return to activities, unlike the long rehabilitation that is required after ACI. A recent study²⁰³ reported that after ACI or MF followed by a long period of rehabilitation, 33% and 26% of sportspeople did not regain full quadriceps power after MF and ACI, respectively. Another study²⁰⁴ reported good results after ACI-P, with 26 of 33 patients having good or excellent results at 10 years, but also noted that patients did not return to full pre-injury activity levels. This may be partly due to the long lay-off during the rehabilitation process. However, those who return to previous activity too early have poorer outcomes than those who wait at least 12 months.²⁰⁵

Conclusion

The evidence base for ACI has improved since the last appraisal by NICE. In most analyses, the ICERs for ACI compared with MF appear to be within a range usually considered acceptable.

The evidence base for ACI is much better than for MF.

Acknowledgements

We thank Peter Auguste for helping with transition probabilities.

We thank the following for advice but absolve them from any responsibility for the final report: Dr Paresh Jobanputra, consultant rheumatologist, Birmingham; Mr Tim Spalding, consultant orthopaedic surgeon, University Hospital Coventry and Warwickshire; and Mr Andrew Sprowson,[†] consultant orthopaedic surgeon, University Hospital Coventry and Warwickshire, and associate professor, Department of Orthopaedics, University of Warwick.

We thank Jill Colquitt and Emma Loveman of Effective Evidence, and Frances Taggart and Tara Gurung, University of Warwick, for data extraction; Dr Michael Crowther, University of Leicester, for statistical advice; and Karoline Munro for secretarial support.

We thank the ACTIVE trial group for permission to report some of their results.

[†]This monograph is dedicated to the memory of Andrew Sprowson, who died in a road traffic accident on 10 April 2015.

Contributions of authors

Hema Mistry reviewed previous studies of the cost-effectiveness, critiqued the cost-effectiveness sections of the manufacturer and Oswestry submissions, created the Warwick model and carried out some of the cost-effectiveness analysis.

Martin Connock carried out the survival analysis in *Chapter 3*.

Joshua Pink carried out other analyses.

Deepson Shyangdan wrote the sections in *Chapters 2* and *5* on trial evidence of clinical effectiveness.

Christine Clar carried out the review of past reviews of clinical effectiveness.

Pamela Royle did literature searches, checked drafts and edited the final report.

Rachel Court extracted data from trials.

Leela C Biant and **Andrew Metcalfe** provided expert orthopaedic advice and commented on drafts.

Norman Waugh wrote *Chapters 1* and *7*, edited drafts of all sections, and edited the final report.

Data sharing statement

There are no further data to share.

References

1. National Institute for Health and Care Excellence (NICE). *Autologous Cartilage Transplantation for Full Thickness Cartilage Defects in Knee Joints*. NICE technology appraisal [TA16]. 2000. URL: www.nice.org.uk/guidance/TA16 (accessed 25 July 2016).
2. Jobanputra P, Parry D, Fry-Smith A, Burls A. Effectiveness of autologous chondrocyte transplantation for hyaline cartilage defects in knees: a rapid and systematic review. *Health Technol Assess* 2001;**5**(11). <https://doi.org/10.3310/hta5110>
3. Clar C, Cummins E, McIntyre L, Thomas S, Lamb J, Bain L, *et al*. Clinical and cost-effectiveness of autologous chondrocyte implantation for cartilage defects in knee joints: systematic review and economic evaluation. *Health Technol Assess* 2005;**9**(47). <https://doi.org/10.3310/hta9470>
4. National Institute for Health and Care Excellence (NICE). *The Use of Autologous Chondrocyte Implantation for the Treatment of Cartilage Defects in the Knee Joints*. NICE technology appraisal guidance [TA89]. 2005. URL: www.nice.org.uk/guidance/ta89 (accessed 25 July 2016).
5. Bentley G, Biant LC, Carrington RW, Akmal M, Goldberg A, Williams AM, *et al*. A prospective, randomised comparison of autologous chondrocyte implantation versus mosaicplasty for osteochondral defects in the knee. *J Bone Joint Surg Br* 2003;**85**:223–30. <https://doi.org/10.1302/0301-620X.85B2.13543>
6. Horas U, Pelinkovic D, Herr G, Aigner T, Schnettler R. Autologous chondrocyte implantation and osteochondral cylinder transplantation in cartilage repair of the knee joint. A prospective, comparative trial. *J Bone Joint Surg Am* 2003;**85A**:185–92. <https://doi.org/10.2106/00004623-200302000-00001>
7. Knutsen G, Engebretsen L, Ludvigsen TC, Drogset JO, Grøntvedt T, Solheim E, *et al*. Autologous chondrocyte implantation compared with microfracture in the knee. A randomized trial. *J Bone Joint Surg Am* 2004;**86A**:455–64. <https://doi.org/10.2106/00004623-200403000-00001>
8. Basad E, Sturz H, Steinmeyer J. [Die Behandlung Chondraler Defekte mit MACI oder Microfracture: erste Ergebnisse einer vergleichenden klinischen Studie.] *Orthopädische Praxis* 2004;**40**:6–10.
9. NICE. *Knee Cartilage Defects – Autologous Chondrocyte Implantation [ID686]*. URL: www.nice.org.uk/guidance/indevelopment/gid-tag446 (accessed 6 January 2017).
10. Indelicato PA, Bittar ES. A perspective of lesions associated with ACL insufficiency of the knee. A review of 100 cases. *Clin Orthop Relat Res* 1985;**198**:77–80.
11. Lewandrowski KU, Müller J, Schollmeier G. Concomitant meniscal and articular cartilage lesions in the femorotibial joint. *Am J Sports Med* 1997;**25**:486–94. <https://doi.org/10.1177/036354659702500411>
12. Mezhev V, Ciccutini FM, Hanna FS, Brennan SL, Wang YY, Urquhart DM, Wluka AE. Does obesity affect knee cartilage? A systematic review of magnetic resonance imaging data. *Obes Rev* 2014;**15**:143–57. <http://dx.doi.org/10.1111/obr.12110>
13. Urquhart DM, Tobing JF, Hanna FS, Berry P, Wluka AE, Ding C, Ciccutini FM. What is the effect of physical activity on the knee joint? A systematic review. *Med Sci Sports Exerc* 2011;**43**:432–42. <http://dx.doi.org/10.1249/MSS.0b013e3181ef5bf8>
14. Arøen A, Løken S, Heir S, Alvik E, Ekeland A, Granlund OG, Engebretsen L. Articular cartilage lesions in 993 consecutive knee arthroscopies. *Am J Sports Med* 2004;**32**:211–15. <https://doi.org/10.1177/0363546503259345>

15. Felson DT. The sources of pain in knee osteoarthritis. *Curr Opin Rheumatol* 2005;**17**:624–8. <https://doi.org/10.1097/01.bor.0000172800.49120.97>
16. Heir S, Nerhus TK, Røtterud JH, Løken S, Ekeland A, Engebretsen L, Arøen A. Focal cartilage defects in the knee impair quality of life as much as severe osteoarthritis: a comparison of knee injury and osteoarthritis outcome score in 4 patient categories scheduled for knee surgery. *Am J Sports Med* 2010;**38**:231–7. <http://dx.doi.org/10.1177/0363546509352157>
17. International Cartilage Repair Society (ICRS). *ICRS Cartilage Injury Evaluation Package*. 2000. URL: www.cartilage.org/_files/contentmanagement/ICRS_evaluation.pdf (accessed 25 July 2016).
18. Moseley JB, O'Malley K, Peterson NJ, Menke TJ, Brody BA, Kuykendall DH, *et al*. A controlled trial of arthroscopic surgery for osteoarthritis of the knee. *N Engl J Med* 2002;**347**:81–8. <http://dx.doi.org/10.1056/NEJMoa013259>
19. Hubbard MJ. Articular debridement versus washout for degeneration of the medial femoral condyle. A five-year study. *J Bone Joint Surg Br* 1996;**78**:217–19.
20. National Institute for Health and Care Excellence (NICE). *Arthroscopic Knee Washout, with or without Debridement, for the Treatment of Osteoarthritis*. NICE interventional procedures guidance [IPG230]. 2007. URL: www.nice.org.uk/guidance/IPG230 (accessed 25 July 2016).
21. Brittberg M, Lindahl A, Nilsson A, Ohlsson C, Isaksson O, Peterson L. Treatment of deep cartilage defects in the knee with autologous chondrocyte transplantation. *N Engl J Med* 1994;**331**:889–95. <http://dx.doi.org/10.1056/NEJM199410063311401>
22. Van der Linden MH, Saris D, Bulstra SK, Buma P. [Treatment of cartilaginous defects in the knee: recommendations from the Dutch Orthopaedic Association.] *Ned Tijdschr Geneesk* 2013;**157**:A5719.
23. Gooding CR, Bartlett W, Bentley G, Skinner JA, Carrington R, Flanagan A. A prospective, randomised study comparing two techniques of autologous chondrocyte implantation for osteochondral defects in the knee: Periosteum covered versus type I/III collagen covered. *Knee* 2006;**13**:203–10. <http://dx.doi.org/10.1016/j.knee.2006.02.011>
24. Gomoll AH, Probst C, Farr J, Cole BJ, Minas T. Use of a type I/III bilayer collagen membrane decreases reoperation rates for symptomatic hypertrophy after autologous chondrocyte implantation. *Am J Sports Med* 2009;**37**(Suppl. 1):20–3. <http://dx.doi.org/10.1177/0363546509348477>
25. Nehrer S, Brix M. Long-term outcomes of chondrocyte-based cartilage repair. *Oper Tech Orthop* 2014;**24**:48–53. <https://doi.org/10.1053/j.oto.2013.12.002>
26. Brittberg M. Cell carriers as the next generation of cell therapy for cartilage repair: a review of the matrix-induced autologous chondrocyte implantation procedure. *Am J Sports Med* 2010;**38**:1259–71. <http://dx.doi.org/10.1177/0363546509346395>
27. Hindle P, Hall AC, Biant LC. Viability of chondrocytes seeded onto a collagen I/III membrane for matrix-induced autologous chondrocyte implantation. *J Orthop Res* 2014;**32**:1495–502. <http://dx.doi.org/10.1002/jor.22701>
28. Mithöfer K, Peterson L, Saris DB, Mandelbaum BR. Evolution and current role of autologous chondrocyte implantation for treatment of articular cartilage defects in the football (soccer) player. *Cartilage* 2012;**3**(Suppl. 1):31–6. <http://dx.doi.org/10.1177/1947603511406532>
29. Gomoll AH, Kamei G, Ochi M, Shetty AA, Zaslav K. Technical enhancements and update on chondrocyte implantation. *Oper Tech Orthop* 2014;**24**:35–47. <https://doi.org/10.1053/j.oto.2014.02.007>

30. Niethammer TR, Pietschmann MF, Horng A, Roßbach BP, Fickscherer A, Jansson V, Müller PE. Graft hypertrophy of matrix-based autologous chondrocyte implantation: a two-year follow-up study of NOVOCART 3D implantation in the knee. *Knee Surg Sports Traumatol Arthrosc* 2014;**22**:1329–36. <http://dx.doi.org/10.1007/s00167-013-2454-7>
31. Harris JD, Siston RA, Brophy RH, Lattermann C, Carey JL, Flanigan DC. Failures, re-operations, and complications after autologous chondrocyte implantation: a systematic review. *Osteoarthr Cartil* 2011;**19**:779–91. <http://dx.doi.org/10.1016/j.joca.2011.02.010>
32. Medical Services Advisory Committee (MSAC). *1140 – Matrix-induced Autologous Chondrocyte Implantation (MACI) and Autologous Chondrocyte Implantation (ACI)*. 2010. URL: www.msac.gov.au/internet/msac/publishing.nsf/Content/app1140-1 (accessed 25 July 2016).
33. Messner K, Gillquist J. Cartilage repair. A critical review. *Acta Orthop Scand* 1996;**67**:523–9. <http://doi.org/10.3109/17453679608996682>
34. National Institute for Health and Care Excellence (NICE). *Mosaicplasty for Knee Cartilage Defects*. NICE interventional procedures guidance [IPG162]. 2006. URL: www.nice.org.uk/guidance/IPG162 (accessed 25 July 2016).
35. Keele University. *ACTIVE Trial Web Site*. 2011. URL: www.active-trial.org.uk/ (accessed 25 July 2016).
36. Messner K, Maletius W. The long-term prognosis for severe damage to weight-bearing cartilage in the knee: a 14-year clinical and radiographic follow-up in 28 young athletes. *Acta Orthop Scand* 1996;**67**:165–8. <https://doi.org/10.3109/17453679608994664>
37. Shelbourne KD, Jari S, Gray T. Outcome of untreated traumatic articular cartilage defects of the knee: a natural history study. *J Bone Joint Surg Am* 2003;**85-A**(Suppl. 2):8–16. <https://doi.org/10.2106/00004623-200300002-00002>
38. Prakash D, Learmonth D. Natural progression of osteo-chondral defect in the femoral condyle. *Knee* 2002;**9**:7–10. [https://doi.org/10.1016/S0968-0160\(01\)00133-8](https://doi.org/10.1016/S0968-0160(01)00133-8)
39. Biant LC, McNicholas MJ, Sprowson AP, Spalding T. The surgical management of symptomatic articular cartilage defects of the knee: consensus statements from United Kingdom knee surgeons. *Knee* 2015;**22**:446–9. <http://dx.doi.org/10.1016/j.knee.2015.06.001>
40. European Medicines Agency (EMA). *ChondroCelect*. 2014. URL: www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/human/medicines/000878/human_med_000698.jsp&mid=WC0b01ac058001d124 (accessed 25 July 2016).
41. PharmaCell. *PharmaCell to Purchase Cell Therapy Production Facility from TiGenix*. 2014. URL: http://pharmacell.nl/press_release_january23/ (accessed 25 July 2016).
42. European Medicines Agency (EMA). *MACI: Matrix-applied Characterised Autologous Cultured Chondrocytes*. 2014. URL: www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/human/medicines/002522/human_med_001660.jsp&mid=WC0b01ac058001d124 (accessed 25 July 2016).
43. Vanlauwe J, Saris DB, Victor J, Almqvist KF, Bellemans J, Luyten FP, TIG/ACT/01/2000&EXT Study Group. Five-year outcome of characterized chondrocyte implantation versus microfracture for symptomatic cartilage defects of the knee: early treatment matters. *Am J Sports Med* 2011;**39**:2566–74. <http://dx.doi.org/10.1177/0363546511422220>
44. Harris JD, Siston RA, Pan XL, Flanigan DC. Autologous chondrocyte implantation a systematic review. *J Bone Joint Surg Am* 2010;**92A**:2220–33. <https://doi.org/10.2106/JBJS.J.00049>
45. Vasiliadis HS, Wasiak J, Salanti G. Autologous chondrocyte implantation for the treatment of cartilage lesions of the knee: a systematic review of randomized studies. *Knee Surg Sports Traumatol Arthrosc* 2010;**18**:1645–55. <https://doi.org/10.1007/s00167-010-1050-3>

46. Vasiliadis HS, Wasiak J. Autologous chondrocyte implantation for full thickness articular cartilage defects of the knee. *Cochrane Database Syst Rev* 2010;**10**:CD003323. <http://dx.doi.org/10.1002/14651858.CD003323.pub3>
47. Goyal D, Goyal A, Keyhani S, Lee EH, Hui JH. Evidence-based status of second- and third-generation autologous chondrocyte implantation over first generation: a systematic review of level I and II studies. *Arthroscopy* 2013;**29**:1872–8. <http://dx.doi.org/10.1016/j.arthro.2013.07.271>
48. Goyal D, Keyhani S, Lee EH, Hui JH. Evidence-based status of microfracture technique: a systematic review of level I and II studies. *Arthroscopy* 2013;**29**:1579–88. <http://dx.doi.org/10.1016/j.arthro.2013.05.027>
49. Naveen S, Robson N, Kamarul T. Comparative analysis of autologous chondrocyte implantation and other treatment modalities: a systematic review. *Eur J Orthop Surg Traumatol* 2012;**22**:89–96. <https://doi.org/10.1007/s00590-011-0798-6>
50. Bekkers JE, Inklaar M, Saris DB. Treatment selection in articular cartilage lesions of the knee: a systematic review. *Am J Sports Med* 2009;**37**(Suppl. 1):148–55. <http://dx.doi.org/10.1177/0363546509351143>
51. Kon E, Filardo G, Di Matteo B, Perdisa F, Marcacci M. Matrix assisted autologous chondrocyte transplantation for cartilage treatment: a systematic review. *Bone Joint Res* 2013;**2**:18–25. <https://doi.org/10.1302/2046-3758.22.2000092>
52. Magnussen RA, Dunn WR, Carey JL, Spindler KP. Treatment of focal articular cartilage defects in the knee: a systematic review. *Clin Orthop Relat Res* 2008;**466**:952–62. <http://dx.doi.org/10.1007/s11999-007-0097-z>
53. Mithöfer K, Hambly K, Della Villa S, Silvers H, Mandelbaum BR. Return to sports participation after articular cartilage repair in the knee: scientific evidence. *Am J Sports Med* 2009;**37**(Suppl. 1):167–76. <https://doi.org/10.1177/0363546509351650>
54. Nakamura N, Miyama T, Engebretsen L, Yoshikawa H, Shino K. Cell-based therapy in articular cartilage lesions of the knee. *Arthroscopy* 2009;**25**:531–52. <http://dx.doi.org/10.1016/j.arthro.2009.02.007>
55. Negrin LL, Vecsei V. Do meta-analyses reveal time-dependent differences between the clinical outcomes achieved by microfracture and autologous chondrocyte implantation in the treatment of cartilage defects of the knee? *J Orthop Sci* 2013;**18**:940–8. <https://doi.org/10.1007/s00776-013-0449-3>
56. Vavken P, Samartzis D. Effectiveness of autologous chondrocyte implantation in cartilage repair of the knee: a systematic review of controlled trials. *Osteoarthritis Cartil* 2010;**18**:857–63. <http://dx.doi.org/10.1016/j.joca.2010.03.005>
57. Bartlett W, Skinner JA, Gooding CR, Carrington RW, Flanagan AM, Briggs TW, Bentley G. Autologous chondrocyte implantation versus matrix-induced autologous chondrocyte implantation for osteochondral defects of the knee: a prospective, randomised study. *J Bone Joint Surg Br* 2005;**87**:640–5. <http://dx.doi.org/10.1302/0301-620X.87B5.15905>
58. Zeifang F, Oberle D, Nierhoff C, Richter W, Moradi B, Schmitt H. Autologous chondrocyte implantation using the original periosteum-cover technique versus matrix-associated autologous chondrocyte implantation: a randomized clinical trial. *Am J Sports Med* 2010;**38**:924–33. <http://dx.doi.org/10.1177/0363546509351499>
59. Erggelet C, Kreuz PC, Mrosek EH, Schagemann JC, Lahm A, Ducommun PP, Ossendorf C. Autologous chondrocyte implantation versus ACI using 3D-bioresorbable graft for the treatment of large full-thickness cartilage lesions of the knee. *Arch Orthop Trauma Surg* 2010;**130**:957–64. <http://dx.doi.org/10.1007/s00402-009-0957-y>

60. Niemeyer P, Pestka JM, Kreuz PC, Ergelet C, Schmal H, Suedkamp NP, Steinwachs M. Characteristic complications after autologous chondrocyte implantation for cartilage defects of the knee joint. *Am J Sports Med* 2008;**36**:2091–9. <http://dx.doi.org/10.1177/0363546508322131>
61. Ferruzzi A, Buda R, Faldini C, Vannini F, Di Caprio F, Luciani D, *et al.* Autologous chondrocyte implantation in the knee joint: open compared with arthroscopic technique. Comparison at a minimum follow-up of five years. *J Bone Joint Surg Am* 2008;**90**(Suppl. 4):90–101. <https://doi.org/10.2106/JBJS.H.00633>
62. Dozin B, Malpeli M, Cancedda R, Bruzzi P, Calcagno S, Molfetta L, *et al.* Comparative evaluation of autologous chondrocyte implantation and mosaicplasty: a multicentered randomized clinical trial. *Clin J Sport Med* 2005;**15**:220–6. <https://doi.org/10.1097/01.jsm.0000171882.66432.80>
63. Horas U, Schnettler R, Pelinkovic D, Herr G, Aigner T. [Osteochondral transplantation versus autogenous chondrocyte transplantation. A prospective comparative clinical study.] *Chirurg* 2000;**71**:1090–7. <https://doi.org/10.1007/s001040051184>
64. Basad E, Ishaque B, Bachmann G, Stürz H, Steinmeyer J. Matrix-induced autologous chondrocyte implantation versus microfracture in the treatment of cartilage defects of the knee: a 2-year randomised study. *Knee Surg Sports Traumatol Arthrosc* 2010;**18**:519–27. <http://dx.doi.org/10.1007/s00167-009-1028-1>
65. Bachmann G, Basad E, Lommel D, Steinmeyer J. [MRI in the follow-up of matrix-supported autologous chondrocyte transplantation (MACI) and microfracture.] *Radiologe* 2004;**44**:773–82. <http://dx.doi.org/10.1007/s00117-004-1084-y>
66. Crawford DC, DeBerardino TM, Williams RJ. NeoCart, an autologous cartilage tissue implant, compared with microfracture for treatment of distal femoral cartilage lesions: an FDA phase-II prospective, randomized clinical trial after two years. *J Bone Joint Surg Am* 2012;**94**:979–89. <http://dx.doi.org/10.2106/JBJS.K.00533>
67. Knutsen G, Drogset JO, Engebretsen L, Grontvedt T, Isaksen V, Ludvigsen TC, *et al.* A randomized trial comparing autologous chondrocyte implantation with microfracture. Findings at five years. *J Bone Joint Surg Am* 2007;**89**:2105–12.
68. Lim HC, Bae JH, Song SH, Park YE, Kim SJ. Current treatments of isolated articular cartilage lesions of the knee achieve similar outcomes. *Clin Orthop Relat Res* 2012;**470**:2261–7. <http://dx.doi.org/10.1007/s11999-012-2304-9>
69. Saris DB, Vanlauwe J, Victor J, Haspl M, Bohnsack M, Fortems Y, *et al.* Characterized chondrocyte implantation results in better structural repair when treating symptomatic cartilage defects of the knee in a randomized controlled trial versus microfracture. *Am J Sports Med* 2008;**36**:235–46. <https://doi.org/10.1177/0363546507311095>
70. Saris DB, Vanlauwe J, Victor J, Almqvist KF, Verdonk R, Bellemans J, *et al.* Treatment of symptomatic cartilage defects of the knee: characterized chondrocyte implantation results in better clinical outcome at 36 months in a randomized trial compared to microfracture. *Am J Sports Med* 2009;**37**(Suppl. 1):10–19. <https://doi.org/10.1177/0363546509350694>
71. Van Assche D, Van Caspel D, Vanlauwe J, Bellemans J, Saris DB, Luyten FP, Staes F. Physical activity levels after characterized chondrocyte implantation versus microfracture in the knee and the relationship to objective functional outcome with 2-year follow-up. *Am J Sports Med* 2009;**37**(Suppl. 1):42–9. <http://dx.doi.org/10.1177/0363546509350296>
72. Van Assche D, Staes F, Van Caspel D, Vanlauwe J, Bellemans J, Saris DB, *et al.* Autologous chondrocyte implantation versus microfracture for knee cartilage injury: a prospective randomized trial, with 2-year follow-up. *Knee Surg Sports Traumatol Arthrosc* 2010;**18**:486–95. <https://doi.org/10.1007/s00167-009-0955-1>

73. Kon E, Gobbi A, Filardo G, Delcogliano M, Zaffagnini S, Marcacci M. Arthroscopic second-generation autologous chondrocyte implantation compared with microfracture for chondral lesions of the knee: prospective nonrandomized study at 5 years. *Am J Sports Med* 2009;**37**:33–41. <http://dx.doi.org/10.1177/0363546508323256>
74. Kon E, Filardo G, Berruto M, Benazzo F, Zanon G, Della Villa S, Marcacci M. Articular cartilage treatment in high-level male soccer players: a prospective comparative study of arthroscopic second-generation autologous chondrocyte implantation versus microfracture. *Am J Sports Med* 2011;**39**:2549–57. <http://dx.doi.org/10.1177/0363546511420688>
75. Minas T, Gomoll AH, Rosenberger R, Royce RO, Bryant T. Increased failure rate of autologous chondrocyte implantation after previous treatment with marrow stimulation techniques. *Am J Sports Med* 2009;**37**:902–8. <http://dx.doi.org/10.1177/0363546508330137>
76. Nejadnik H, Hui JH, Feng Choong EP, Tai BC, Lee EH. Autologous bone marrow-derived mesenchymal stem cells versus autologous chondrocyte implantation: an observational cohort study. *Am J Sports Med* 2010;**38**:1110–16. <http://dx.doi.org/10.1177/0363546509359067>
77. Visna P, Pasa L, Cizmár I, Hart R, Hoch J. Treatment of deep cartilage defects of the knee using autologous chondrograft transplantation and by abrasive techniques – a randomized controlled study. *Acta Chir Belg* 2004;**104**:709–14. <https://doi.org/10.1080/00015458.2004.11679648>
78. Bentley G, Biant LC, Vijayan S, Macmull S, Skinner JA, Carrington RW. Minimum ten-year results of a prospective randomised study of autologous chondrocyte implantation versus mosaicplasty for symptomatic articular cartilage lesions of the knee. *J Bone Joint Surg Br* 2012;**94**:504–9. <http://dx.doi.org/10.1302/0301-620X.94B4.27495>
79. Biant LC, Bentley G, Vijayan S, Skinner JA, Carrington RW. Long-term results of autologous chondrocyte implantation in the knee for chronic chondral and osteochondral defects. *Am J Sports Med* 2014;**42**:2178–83. <http://dx.doi.org/10.1177/0363546514539345>
80. Nawaz SZ, Bentley G, Briggs TW, Carrington RW, Skinner JA, Gallagher KR, Dhinsa BS. Autologous chondrocyte implantation in the knee: mid-term to long-term results. *J Bone Joint Surg Am* 2014;**96**:824–30. <http://dx.doi.org/10.2106/JBJS.L.01695>
81. Pestka JM, Bode G, Salzmann G, Südkamp NP, Niemeyer P. Clinical outcome of autologous chondrocyte implantation for failed microfracture treatment of full-thickness cartilage defects of the knee joint. *Am J Sports Med* 2012;**40**:325–31. <http://dx.doi.org/10.1177/0363546511425651>
82. Mithöfer K, McAdams T, Williams RJ, Kreuz PC, Mandelbaum BR. Clinical efficacy of the microfracture technique for articular cartilage repair in the knee: an evidence-based systematic analysis. *Am J Sports Med* 2009;**37**:2053–63. <http://dx.doi.org/10.1177/0363546508328414>
83. Behery OA, Harris JD, Karnes JM, Siston RA, Flanigan DC. Factors influencing the outcome of autologous chondrocyte implantation: a systematic review. *J Knee Surg* 2013;**26**:203–11. <http://dx.doi.org/10.1055/s-0032-1329231>
84. Behery O, Siston RA, Harris JD, Flanigan DC. Treatment of cartilage defects of the knee: expanding on the existing algorithm. *Clin J Sport Med* 2014;**24**:21–30. <http://dx.doi.org/10.1097/JSM.0000000000000004>
85. Chalmers PN, Vigneswaran H, Harris JD, Cole BJ. Activity-related outcomes of articular cartilage surgery: a systematic review. *Cartilage* 2013;**4**:193–203. <http://dx.doi.org/10.1177/1947603513481603>
86. Howard JS, Lattermann C, Hoch JM, Mattacola CG, Medina McKeon JM. Comparing responsiveness of six common patient-reported outcomes to changes following autologous chondrocyte implantation: a systematic review and meta-analysis of prospective studies. *Cartilage* 2013;**4**:97–110. <http://dx.doi.org/10.1177/1947603512470684>

87. Bartlett W, Gooding CR, Carrington RW, Skinner JA, Briggs TW, Bentley G. Autologous chondrocyte implantation at the knee using a bilayer collagen membrane with bone graft. A preliminary report. *J Bone Joint Surg Br* 2005;**87**:330–2. <https://doi.org/10.1302/0301-620X.87B3.15552>
88. Benthien JP, Schwaninger M, Behrens P. We do not have evidence based methods for the treatment of cartilage defects in the knee. *Knee Surg Sports Traumatol Arthrosc* 2011;**19**:543–52. <http://dx.doi.org/10.1007/s00167-010-1271-5>
89. Cole BJ, Farr J, Winalski CS, Hosea T, Richmond J, Mandelbaum B, De Deyne PG. Outcomes after a single-stage procedure for cell-based cartilage repair: a prospective clinical safety trial with 2-year follow-up. *Am J Sports Med* 2011;**39**:1170–9. <http://dx.doi.org/10.1177/0363546511399382>
90. Ebert JR, Lloyd DG, Ackland T, Wood DJ. Knee biomechanics during walking gait following matrix-induced autologous chondrocyte implantation. *Clin Biomech* 2010;**25**:1011–17. <http://dx.doi.org/10.1016/j.clinbiomech.2010.07.004>
91. Ebert JR, Fallon M, Zheng MH, Wood DJ, Ackland TR. A randomized trial comparing accelerated and traditional approaches to postoperative weightbearing rehabilitation after matrix-induced autologous chondrocyte implantation: findings at 5 years. *Am J Sports Med* 2012;**40**:1527–37. <http://dx.doi.org/10.1177/0363546512445167>
92. Edwards PK, Ackland TR, Ebert JR. Accelerated weightbearing rehabilitation after matrix-induced autologous chondrocyte implantation in the tibiofemoral joint: early clinical and radiological outcomes. *Am J Sports Med* 2013;**41**:2314–24. <http://dx.doi.org/10.1177/0363546513495637>
93. Harris JD, Brophy RH, Siston RA, Flanigan DC. Treatment of chondral defects in the athlete's knee. *Arthroscopy* 2010;**26**:841–52. <https://doi.org/10.1016/j.arthro.2009.12.030>
94. Panseri S, Russo A, Cunha C, Bondi A, Di Martino A, Patella S, *et al.* Osteochondral tissue engineering approaches for articular cartilage and subchondral bone regeneration. *Knee Surg Sports Traumatol Arthrosc* 2012;**20**:1182–91. <https://doi.org/10.1007/s00167-011-1655-1>
95. Rodriguez-Merchan EC. The treatment of cartilage defects in the knee joint: microfracture, mosaicplasty, and autologous chondrocyte implantation. *Am J Orthop (Belle Mead NJ)* 2012;**41**:236–9.
96. Ruano-Ravina A, Jato Díaz M. Autologous chondrocyte implantation: a systematic review. *Osteoarthr Cartil* 2006;**14**:47–51. <http://dx.doi.org/10.1016/j.joca.2005.07.017>
97. Toonstra JL, Howard JS, Uhl TL, English RA, Mattacola CG. The role of rehabilitation following autologous chondrocyte implantation: a retrospective chart review. *Int J Sports Phys Ther* 2013;**8**:670–9.
98. Trinh TQ, Harris JD, Siston RA, Flanigan DC. Improved outcomes with combined autologous chondrocyte implantation and patellofemoral osteotomy versus isolated autologous chondrocyte implantation. *Arthroscopy* 2013;**29**:566–74. <http://dx.doi.org/10.1016/j.arthro.2012.10.008>
99. United Healthcare. *Medical Policy: Autologous Chondrocyte Transplantation in the Knee*. 2013. URL: [www.unitedhealthcareonline.com/ccmcontent/ProviderII/UHC/en-US/Assets/ProviderStaticFiles/ProviderStaticFilesPdf/Tools and Resources/Policies and Protocols/Medical Policies/Medical Policies/Autologous_Chondrocyte_Transplantation_in_the_Knee.pdf](http://www.unitedhealthcareonline.com/ccmcontent/ProviderII/UHC/en-US/Assets/ProviderStaticFiles/ProviderStaticFilesPdf/Tools%20and%20Resources/Policies%20and%20Protocols/Medical%20Policies/Medical%20Policies/Autologous_Chondrocyte_Transplantation_in_the_Knee.pdf) (accessed 25 July 2016).
100. Saris D, Price A, Widuchowski W, Bertrand-Marchand M, Caron J, Drogset JO, *et al.* Matrix-applied characterized autologous cultured chondrocytes versus microfracture: two-year follow-up of a prospective randomized trial. *Am J Sports Med* 2014;**42**:1384–94. <http://dx.doi.org/10.1177/0363546514528093>

101. Steadman JR, Rodkey WG, Rodrigo JJ. Microfracture: surgical technique and rehabilitation to treat chondral defects. *Clin Orthop Relat Res* 2001;**391**(Suppl.):362–9. <https://doi.org/10.1097/00003086-200110001-00033>
102. Niemeyer P, Salzmann G, Feucht M, Pestka J, Porichis S, Ogon P, *et al.* First-generation versus second-generation autologous chondrocyte implantation for treatment of cartilage defects of the knee: a matched-pair analysis on long-term clinical outcome. *Int Orthop* 2014;**38**:2065–70. <http://dx.doi.org/10.1007/s00264-014-2368-0>
103. Ebert JR, Robertson WB, Woodhouse J, Fallon M, Zheng MH, Ackland T, Wood DJ. Clinical and magnetic resonance imaging-based outcomes to 5 years after matrix-induced autologous chondrocyte implantation to address articular cartilage defects in the knee. *Am J Sports Med* 2011;**39**:753–63. <http://dx.doi.org/10.1177/0363546510390476>
104. Negrin L, Kutscha-Lissberg F, Gartlehner G, Vecsei V. Clinical outcome after microfracture of the knee: a meta-analysis of before/after-data of controlled studies. *Int Orthop* 2012;**36**:43–50. <http://dx.doi.org/10.1007/s00264-011-1364-x>
105. NIH National Heart Lung and Blood Institute. *Quality Assessment Tool for Case Series Studies*. 2014. URL: www.nhlbi.nih.gov/health-pro/guidelines/in-develop/cardiovascular-risk-reduction/tools/case_series (accessed 25 July 2016).
106. Asik M, Ciftci F, Sen C, Erdil M, Atalar A. The microfracture technique for the treatment of full-thickness articular cartilage lesions of the knee: midterm results. *Arthroscopy* 2008;**24**:1214–20. <http://dx.doi.org/10.1016/j.arthro.2008.06.015>
107. Beris AE, Lykissas MG, Kostas-Agnantis I, Manoudis GN. Treatment of full-thickness chondral defects of the knee with autologous chondrocyte implantation: a functional evaluation with long-term follow-up. *Am J Sports Med* 2012;**40**:562–7. <https://doi.org/10.1177/0363546511428778>
108. Bhosale AM, Kuiper JH, Johnson WE, Harrison PE, Richardson JB. Midterm to long-term longitudinal outcome of autologous chondrocyte implantation in the knee joint: a multilevel analysis. *Am J Sports Med* 2009;**37**(Suppl. 1):131–8. <http://dx.doi.org/10.1177/0363546509350555>
109. Browne JE, Anderson AF, Arciero R, Mandelbaum B, Moseley JB, Micheli LJ, *et al.* Clinical outcome of autologous chondrocyte implantation at 5 years in US subjects. *Clin Orthop Relat Res* 2005;**436**:237–45. <https://doi.org/10.1097/00003086-200507000-00036>
110. Gomoll AH, Gillogly SD, Cole BJ, Farr J, Arnold R, Hussey K, Minas T. Autologous chondrocyte implantation in the patella: a multicenter experience. *Am J Sports Med* 2014;**42**:1074–81. <http://dx.doi.org/10.1177/0363546514523927>
111. Jungmann PM, Salzmann GM, Schmal H, Pestka JM, Südkamp NP, Niemeyer P. Autologous chondrocyte implantation for treatment of cartilage defects of the knee: what predicts the need for reintervention? *Am J Sports Med* 2012;**40**:58–67. <http://dx.doi.org/10.1177/0363546511423522>
112. Krych AJ, Harnly HW, Rodeo SA, Williams RJ III. Activity levels are higher after osteochondral autograft transfer mosaicplasty than after microfracture for articular cartilage defects of the knee: a retrospective comparative study. *J Bone Joint Surg Am* 2012;**94**:971–8. <https://doi.org/10.2106/JBJS.K.00815>
113. Layton A, Arnold RJ, Graham J, Frasco MA, Cote E, Lynch NM. Long-term failure rates associated with knee microfracture surgery. *Value Health* 2015;**18**:A156. <https://doi.org/10.1016/j.jval.2015.03.905>
114. Layton A, Arnold RJ, Graham J, Frasco MA, Cote E, Lynch NM. *Long-term Failure Rates Associated with Knee Microfracture Surgery [poster]*. 2015. URL: http://quorumconsulting.com/docs/MicrofractureFailureRates_ISPOR_May2015.pdf (accessed 25 July 2016).

115. Moseley JB, Anderson AF, Browne JE, Mandelbaum BR, Micheli LJ, Fu F, Erggelet C. Long-term durability of autologous chondrocyte implantation: a multicenter, observational study in US patients. *Am J Sports Med* 2010;**38**:238–46. <http://dx.doi.org/10.1177/0363546509348000>
116. Rogers BA, David LA, Briggs TW. Sequential outcome following autologous chondrocyte implantation of the knee: a six-year follow-up. *Int Orthop* 2010;**34**:959–64. <http://dx.doi.org/10.1007/s00264-009-0842-x>
117. Niemeyer P, Porichis S, Steinwachs M, Erggelet C, Kreuz PC, Schmal H, et al. Long-term outcomes after first-generation autologous chondrocyte implantation for cartilage defects of the knee. *Am J Sports Med* 2014;**42**:150–7. <http://dx.doi.org/10.1177/0363546513506593>
118. Peterson L, Vasiliadis HS, Brittberg M, Lindahl A. Autologous chondrocyte implantation: a long-term follow-up. *Am J Sports Med* 2010;**38**:1117–24. <http://dx.doi.org/10.1177/0363546509357915>
119. Peterson L, Brittberg M, Kiviranta I, Akerlund EL, Lindahl A. Autologous chondrocyte transplantation. Biomechanics and long-term durability. *Am J Sports Med* 2002;**30**:2–12.
120. Peterson L, Minas T, Brittberg M, Lindahl A. Treatment of osteochondritis dissecans of the knee with autologous chondrocyte transplantation: results at two to ten years. *J Bone Joint Surg Am* 2003;**85**(Suppl. 2):17–24. <https://doi.org/10.2106/00004623-200300002-00003>
121. Salzmann GM, Sah B, Südkamp NP, Niemeyer P. Clinical outcome following the first-line, single lesion microfracture at the knee joint. *Arch Orthop Trauma Surg* 2013;**133**:303–10. <http://dx.doi.org/10.1007/s00402-012-1660-y>
122. Shive MS, Stanish WD, McCormack R, Forriol F, Mohtadi N, Pelet S, et al. BST-CarGel® Treatment maintains cartilage repair superiority over microfracture at 5 years in a multicenter randomized controlled trial. *Cartilage* 2015;**6**:62–72. <https://doi.org/10.1177/1947603514562064>
123. Frappier J, Stanish W, Brittberg M, Steinwachs M, Crowe L, Castelo D, Restrepo A. Economic evaluation of BST-CarGel as an adjunct to microfracture vs microfracture alone in knee cartilage surgery. *J Med Econ* 2014;**17**:266–78. <http://dx.doi.org/10.3111/13696998.2014.897626>
124. Solheim E, Hegna J, Inderhaug E, Øyen J, Harlem T, Strand T. Results at 10-14 years after microfracture treatment of articular cartilage defects in the knee. *Knee Surg Sports Traumatol Arthrosc* 2016;**24**:1587–93. <http://dx.doi.org/10.1007/s00167-014-3443-1>
125. Solheim E, Oyen J, Hegna J, Austgulen OK, Harlem T, Strand T. Microfracture treatment of single or multiple articular cartilage defects of the knee: a 5-year median follow-up of 110 patients. *Knee Surg Sports Traumatol Arthrosc* 2010;**18**:504–8. <https://doi.org/10.1007/s00167-009-0974-y>
126. Steadman JR, Briggs KK, Rodrigo JJ, Kocher MS, Gill TJ, Rodkey WG. Outcomes of microfracture for traumatic chondral defects of the knee: average 11-year follow-up. *Arthroscopy* 2003;**19**:477–84. <http://dx.doi.org/10.1053/jars.2003.50112>
127. Gobbi A, Karnatzikos G, Kumar A. Long-term results after microfracture treatment for full-thickness knee chondral lesions in athletes. *Knee Surg Sports Traumatol Arthrosc* 2014;**22**:1986–96. <https://doi.org/10.1007/s00167-013-2676-8>
128. Gudas R, Gudaite A, Pocius A, Gudiene A, Cekanauskas E, Monastyreckiene E, Basevicius A. Ten-year follow-up of a prospective, randomized clinical study of mosaic osteochondral autologous transplantation versus microfracture for the treatment of osteochondral defects in the knee joint of athletes. *Am J Sports Med* 2012;**40**:2499–508. <http://dx.doi.org/10.1177/0363546512458763>
129. Bert JM. Abandoning microfracture of the knee: has the time come? *Arthroscopy* 2015;**31**:501–5. <http://dx.doi.org/10.1016/j.arthro.2014.12.018>

130. Gudas R, Gudaite A, Mickevicius T, Masiulis N, Simonaityte R, Cekanauskas E, *et al.* Comparison of osteochondral autologous transplantation, microfracture, or debridement techniques in articular cartilage lesions associated with anterior cruciate ligament injury: a prospective study with a 3-year follow-up. *Arthroscopy* 2013;**29**:89–97. <https://doi.org/10.1016/j.arthro.2012.06.009>
131. Guyot P, Ades AE, Ouwens MJ, Welton NJ. Enhanced secondary analysis of survival data: reconstructing the data from published Kaplan-Meier survival curves. *BMC Med Res Methodol* 2012;**12**:9. <http://dx.doi.org/10.1186/1471-2288-12-9>
132. Crowther MJ, Lambert PC. *stgenreg*: a Stata package for general parametric survival analysis. *J Stat Softw* 2013;**53**:1–17. <https://doi.org/10.18637/jss.v053.i12>
133. Collet D. *Modelling Survival Data in Medical Research*. 2nd edn. Boca Raton, FL: Chapman & Hall; 2003.
134. Bae DK, Song SJ, Yoon KH, Heo DB, Kim TJ. Survival analysis of microfracture in the osteoarthritic knee-minimum 10-year follow-up. *Arthroscopy* 2013;**29**:244–50. <http://dx.doi.org/10.1016/j.arthro.2012.09.006>
135. Machin D, Cheung Y, Parmar M. *Survival Analysis: A Practical Approach*. Chichester: John Wiley & Sons; 2006. <https://doi.org/10.1002/0470034572>
136. Minas T, Von Keudell A, Bryant T, Gomoll AH. The John Insall award: a minimum 10-year outcome study of autologous chondrocyte implantation. *Clin Orthop Relat Res* 2014;**472**:41–51. <https://doi.org/10.1007/s11999-013-3146-9>
137. Salzmann GM, Sah B, Südkamp NP, Niemeyer P. Reoperative characteristics after microfracture of knee cartilage lesions in 454 patients. *Knee Surg Sports Traumatol Arthrosc* 2013;**21**:365–71. <http://dx.doi.org/10.1007/s00167-012-1973-y>
138. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, *et al.* Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement. *Value Health* 2013;**16**:e1–5. <https://doi.org/10.1016/j.jval.2013.02.010>
139. Philips Z, Ginnelly L, Sculpher M, Claxton K, Golder S, Riemsma R, *et al.* Review of guidelines for good practice in decision-analytic modelling in health technology assessment. *Health Technol Assess* 2004;**8**(36). <https://doi.org/10.3310/hta8360>
140. Derrett S, Stokes EA, James M, Bartlett W, Bentley G. Cost and health status analysis after autologous chondrocyte implantation and mosaicplasty: a retrospective comparison. *Int J Technol Assess Health Care* 2005;**21**:359–67. <https://doi.org/10.1017/S0266462305050476>
141. Gerlier L, Lamotte M, Wille M, Kreuz PC, Vanlauwe J, Dubois D, Meurgey FM. The cost utility of autologous chondrocytes implantation using ChondroCelect® in symptomatic knee cartilage lesions in Belgium. *Pharmacoeconomics* 2010;**28**:1129–46. <http://dx.doi.org/10.2165/11584920-000000000-00000>
142. Samuelson EM, Brown DE. Cost-effectiveness analysis of autologous chondrocyte implantation: a comparison of periosteal patch versus type VIII collagen membrane. *Am J Sports Med* 2012;**40**:1252–8. <http://dx.doi.org/10.1177/0363546512441586>
143. Koerber F, Rolauffs B, Rogowski W. Early evaluation and value-based pricing of regenerative medicine technologies. *Regen Med* 2013;**8**:747–58. <http://dx.doi.org/10.2217/rme.13.69>
144. Medical Services Advisory Committee (MSAC). 1273 – *Consultation Decision Analytic Protocol (DAP) to Guide the Assessment of Matrix-induced Autologous Chondrocyte Implantation (MACI)*. 2012. URL: www.msac.gov.au/. (accessed 25 July 2016).

145. Kunzl M, Wild C, Mathis S, Johansson T. *Autologous Chondrocyte Implantation*. Ludwig Boltzmann Institut für Health Technology Assessment; 2009. URL: <http://eprints.hta.lbg.ac.at/865/> (accessed 25 July 2016).
146. Filardo G, Kon E, Andriolo L, Vannini F, Buda R, Ferruzzi A, *et al*. Does patient sex influence cartilage surgery outcome? Analysis of results at 5-year follow-up in a large cohort of patients treated with matrix-assisted autologous chondrocyte transplantation. *Am J Sports Med* 2013;**41**:1827–34. <https://doi.org/10.1177/0363546513480780>
147. Forster MC. Survival analysis of primary cemented total knee arthroplasty: which designs last? *J Arthroplasty* 2003;**18**:265–70. <http://dx.doi.org/10.1054/arth.2003.50051>
148. Bernard J, Lemon M, Patterson MH. Arthroscopic washout of the knee: a 5-year survival analysis. *Knee* 2004;**11**:233–5. [http://dx.doi.org/10.1016/S0968-0160\(03\)00108-X](http://dx.doi.org/10.1016/S0968-0160(03)00108-X)
149. NHS Choices. *Knee Replacement*. 2014. URL: www.nhs.uk/Conditions/Knee-replacement/Pages/Kneereplacementexplained.aspx (accessed 25 July 2016).
150. Mahomed NN, Barrett J, Katz JN, Baron JA, Wright J, Losina E. Epidemiology of total knee replacement in the United States Medicare population. *J Bone Joint Surg Am* 2005;**87**:1222–8. <http://dx.doi.org/10.2106/JBJS.D.02546>
151. Department of Health (DH). *NHS Reference Costs 2012 to 2013*. London: DH; 2013. URL: www.gov.uk/government/publications/nhs-reference-costs-2012-to-2013 (accessed 25 July 2016).
152. Kind P, Hardman G, Macran S. *UK Population Norms for EQ-5D*. Discussion Paper 172. 2009. URL: [www.york.ac.uk/media/che/documents/papers/discussionpapers/CHE Discussion Paper 172.pdf](http://www.york.ac.uk/media/che/documents/papers/discussionpapers/CHE%20Discussion%20Paper%20172.pdf) (accessed 25 July 2016).
153. Vijayan S, Bentley G, Rahman J, Briggs TW, Skinner JA, Carrington RW. Revision cartilage cell transplantation for failed autologous chondrocyte transplantation in chronic osteochondral defects of the knee. *Bone Joint J* 2014;**96-B**:54–8. <http://dx.doi.org/10.1302/0301-620X.96B1.31979>
154. Curtis L. *Unit Costs of Health and Social Care*. Canterbury: PSSRU, University of Kent; 2013. URL: www.pssru.ac.uk/project-pages/unit-costs/2013/ (accessed 25 July 2016).
155. McCarthy HS, Roberts S. A histological comparison of the repair tissue formed when using either Chondrogide® or periosteum during autologous chondrocyte implantation. *Osteoarthritis Cartilage* 2013;**21**:2048–57. <https://doi.org/10.1016/j.joca.2013.10.004>
156. Filardo G, Kon E, Di Martino A, Iacono F, Marcacci M. Arthroscopic second-generation autologous chondrocyte implantation: a prospective 7-year follow-up study. *Am J Sports Med* 2011;**39**:2153–60. <http://dx.doi.org/10.1177/0363546511415658>
157. National Joint Registry. *11th Annual Report: Surgical Data to 31 December 2013. Part 3: 3.3 Outcomes after Primary Knee Replacement*. 2014. URL: www.njrcentre.org.uk/njrcentre/Reports,PublicationsandMinutes/Annualreports/tabid/86/Default.aspx (accessed 25 July 2016).
158. Dong H, Buxton M. Early assessment of the likely cost-effectiveness of a new technology: a Markov model with probabilistic sensitivity analysis of computer-assisted total knee replacement. *Int J Technol Assess Health Care* 2006;**22**:191–202. <http://dx.doi.org/10.1017/S0266462306051014>
159. Jansson KA, Granath F. Health-related quality of life (EQ-5D) before and after orthopedic surgery. *Acta Orthop* 2011;**82**:82–9. <https://doi.org/10.3109/17453674.2010.548026>
160. Office for National Statistics (ONS). *National Life Tables, United Kingdom, 2011–2013*. 2014. URL: www.ons.gov.uk/ons/dcp171778_377972.pdf (accessed 25 July 2016).
161. Briggs A, Claxton K, Sculpher M. *Decision Modelling for Health Economic Evaluation*. Oxford: Oxford University Press; 2006.

162. Fox M, Mealing S, Anderson R, Dean J, Stein K, Price A, *et al.* The clinical effectiveness and cost-effectiveness of cardiac resynchronisation (biventricular pacing) for heart failure: systematic review and economic model. *Health Technol Assess* 2007;**11**(47). <https://doi.org/10.3310/hta11470>
163. Drummond M, McGuire A. *Economic Evaluation in HealthCare: Merging Theory with Practice*. Oxford: Oxford University Press; 2001.
164. Beard D, Price A, Cook J, Fitzpatrick R, Carr A, Campbell M, *et al.* Total or Partial Knee Arthroplasty Trial - TOPKAT: study protocol for a randomised controlled trial. *Trials* 2013;**14**:292. <http://dx.doi.org/10.1186/1745-6215-14-292>
165. Harris JD, Erickson BJ, Abrams GD, Cvetanovich GL, McCormick FM, Gupta AK, *et al.* Methodologic quality of knee articular cartilage studies. *Arthroscopy* 2013;**29**:1243–52.e5. <http://dx.doi.org/10.1016/j.arthro.2013.02.023>
166. De Windt TS, Vonk LA, Brittberg M, Saris DB. Treatment and prevention of (early) osteoarthritis using articular cartilage repair-fact or fiction? A systematic review. *Cartilage* 2013;**4**(Suppl. 3):5–12. <http://dx.doi.org/10.1177/1947603513486560>
167. Creamer P, Lethbridge-Cejku M, Costa P, Tobin JD, Herbst JH, Hochberg MC. The relationship of anxiety and depression with self-reported knee pain in the community: data from the Baltimore Longitudinal Study of Aging. *Arthritis Care Res* 1999;**12**:3–7. [https://doi.org/10.1002/1529-0131\(199902\)12:1<3::AID-ART2>3.0.CO;2-K](https://doi.org/10.1002/1529-0131(199902)12:1<3::AID-ART2>3.0.CO;2-K)
168. Dieppe PA, Lohmander LS. Pathogenesis and management of pain in osteoarthritis. *Lancet* 2005;**365**:965–73. [http://dx.doi.org/10.1016/S0140-6736\(05\)71086-2](http://dx.doi.org/10.1016/S0140-6736(05)71086-2)
169. Zhang W, Doherty M, Peat G, Bierma-Zeinstra MA, Arden NK, Bresnihan B, *et al.* EULAR evidence-based recommendations for the diagnosis of knee osteoarthritis. *Ann Rheum Dis* 2010;**69**:483–9. <http://dx.doi.org/10.1136/ard.2009.113100>
170. Pereira D, Severo M, Barros H, Branco J, Santos RA, Ramos E. The effect of depressive symptoms on the association between radiographic osteoarthritis and knee pain: a cross-sectional study. *BMC Musculoskelet Disord* 2013;**14**:214. <http://dx.doi.org/10.1186/1471-2474-14-214>
171. Bedson J, Croft PR. The discordance between clinical and radiographic knee osteoarthritis: a systematic search and summary of the literature. *BMC Musculoskelet Disord* 2008;**9**:116. <http://dx.doi.org/10.1186/1471-2474-9-116>
172. Link TM. Correlations between joint morphology and pain and between magnetic resonance imaging, histology, and micro-computed tomography. *J Bone Joint Surg Am* 2009;**91**(Suppl. 1):30–2. <http://dx.doi.org/10.2106/JBJS.H.01313>
173. Neogi T, Felson D, Niu J, Nevitt M, Lewis CE, Aliabadi P, *et al.* Association between radiographic features of knee osteoarthritis and pain: results from two cohort studies. *BMJ* 2009;**339**:b2844. <http://dx.doi.org/10.1136/bmj.b2844>
174. Kellgren JH, Lawrence JS. Radiological assessment of osteo-arthrosis. *Ann Rheum Dis* 1957;**16**:494–502. <https://doi.org/10.1136/ard.16.4.494>
175. Link TM, Steinbach LS, Ghosh S, Ries M, Lu Y, Lane N, Majumdar S. Osteoarthritis: MR imaging findings in different stages of disease and correlation with clinical findings. *Radiology* 2003;**226**:373–81. <http://dx.doi.org/10.1148/radiol.2262012190>
176. Williams A, Sharma L, McKenzie CA, Prasad PV, Burstein D. Delayed gadolinium-enhanced magnetic resonance imaging of cartilage in knee osteoarthritis: findings at different radiographic stages of disease and relationship to malalignment. *Arthritis Rheum* 2005;**52**:3528–35. <https://doi.org/10.1002/art.21388>

177. Kijowski R, Blankenbaker D, Stanton P, Fine J, De Smet A. Arthroscopic validation of radiographic grading scales of osteoarthritis of the tibiofemoral joint. *AJR Am J Roentgenol* 2006;**187**:794–9. <http://dx.doi.org/10.2214/AJR.05.1123>
178. Szebenyi B, Hollander AP, Dieppe P, Quilty B, Duddy J, Clarke S, Kirwan JR. Associations between pain, function, and radiographic features in osteoarthritis of the knee. *Arthritis Rheum* 2006;**54**:230–5. <https://doi.org/10.1002/art.21534>
179. Wright RW. Osteoarthritis classification scales: interobserver reliability and arthroscopic correlation. *J Bone Joint Surg Am* 2014;**96**:1145–51. <https://doi.org/10.2106/JBJS.M.00929>
180. Sheehy L, Culham E, McLean L, Niu J, Lynch J, Segal NA, *et al.* Validity and sensitivity to change of three scales for the radiographic assessment of knee osteoarthritis using images from the Multicenter Osteoarthritis Study (MOST). *Osteoarthr Cartil* 2015;**23**:1491–8. <http://dx.doi.org/10.1016/j.joca.2015.05.003>
181. Schiphof D, Boers M, Bierma-Zeinstra SM. Differences in descriptions of Kellgren and Lawrence grades of knee osteoarthritis. *Ann Rheum Dis* 2008;**67**:1034–6. <http://dx.doi.org/10.1136/ard.2007.079020>
182. Felson DT, Niu J, Guermazi A, Sack B, Aliabadi P. Defining radiographic incidence and progression of knee osteoarthritis: suggested modifications of the Kellgren and Lawrence scale. *Ann Rheum Dis* 2011;**70**:1884–6. <http://dx.doi.org/10.1136/ard.2011.155119>
183. Thorstensson CA, Andersson ML, Jönsson H, Saxne T, Petersson IF. Natural course of knee osteoarthritis in middle-aged subjects with knee pain: 12-year follow-up using clinical and radiographic criteria. *Ann Rheum Dis* 2009;**68**:1890–3. <http://dx.doi.org/10.1136/ard.2008.095158>
184. Leyland KM, Hart DJ, Javaid MK, Judge A, Kiran A, Soni A, *et al.* The natural history of radiographic knee osteoarthritis: a fourteen-year population-based cohort study. *Arthritis Rheum* 2012;**64**:2243–51. <http://dx.doi.org/10.1002/art.34415>
185. Minas T, Gomoll AH, Solhpour S, Rosenberger R, Probst C, Bryant T. Autologous chondrocyte implantation for joint preservation in patients with early osteoarthritis. *Clin Orthop Relat Res* 2010;**468**:147–57. <https://doi.org/10.1007/s11999-009-0998-0>
186. Jaiswal PK, Bentley G, Carrington RW, Skinner JA, Briggs TW. The adverse effect of elevated body mass index on outcome after autologous chondrocyte implantation. *J Bone Joint Surg Br* 2012;**94**:1377–81. <http://dx.doi.org/10.1302/0301-620X.94B10.29388>
187. Nakamura N, Hui J, Koizumi K, Yasui Y, Nishii Y, Lad D, *et al.* Stem cell therapy in cartilage repair: culture-free and cell culture-based methods. *Oper Tech Orthop* 2014;**24**:54–60. <https://doi.org/10.1053/j.oto.2014.02.006>
188. Perera JR, Gikas PD, Bentley G. The present state of treatments for articular cartilage defects in the knee. *Ann R Coll Surg Engl* 2012;**94**:381–7. <http://dx.doi.org/10.1308/003588412X13171221592573>
189. Filardo G, Kon E, Roffi A, Di Martino A, Marcacci M. Scaffold-based repair for cartilage healing: a systematic review and technical note. *Arthroscopy* 2013;**29**:174–86. <http://dx.doi.org/10.1016/j.arthro.2012.05.891>
190. Buda R, Vannini F, Cavallo M, Grigolo B, Cenacchi A, Giannini S. Osteochondral lesions of the knee: a new one-step repair technique with bone-marrow-derived cells. *J Bone Joint Surg Am* 2010;**92**(Suppl. 2):2–11. <http://dx.doi.org/10.2106/JBJS.J.00813>
191. The Robert Jones and Agnes Hunt Orthopaedic Hospital NHS Foundation Trust. ASCOT: *Autologous Stem Cells, Chondrocytes or the Two?* 2014. URL: <http://public.ukcrn.org.uk/Search/StudyDetail.aspx?StudyID=12383> (accessed 25 July 2016).

192. Hendricks J, Verdonk P, Widuchowski J, Snow M, Weiss W, Kruczynski J, et al. *First Clinical Experience with INSTRUMENT: a Single Surgery, Autologous Cell Based Technology for Cartilage Repair*. 2014. URL: [www.cellcotec.com/content/P187_first_clinical_experience_with INSTRUMENT final2.pdf](http://www.cellcotec.com/content/P187_first_clinical_experience_with_INSTRUMENT_final2.pdf) (accessed 25 July 2016).
193. Adachi N, Ochi M, Deie M, Nakamae A, Kamei G, Uchio Y, Iwasa J. Implantation of tissue-engineered cartilage-like tissue for the treatment for full-thickness cartilage defects of the knee. *Knee Surg Sports Traumatol Arthrosc* 2014;**22**:1241–8. <http://dx.doi.org/10.1007/s00167-013-2521-0>
194. Choi NY, Kim BW, Yeo WJ, Kim HB, Suh DS, Kim JS, et al. Gel-type autologous chondrocyte (Chondron) implantation for treatment of articular cartilage defects of the knee. *BMC Musculoskelet Disord* 2010;**11**:103. <http://dx.doi.org/10.1186/1471-2474-11-103>
195. Mizuno M, Kobayashi S, Takebe T, Kan H, Yabuki Y, Matsuzaki T, et al. Brief report: reconstruction of joint hyaline cartilage by autologous progenitor cells derived from ear elastic cartilage. *Stem Cells* 2014;**32**:816–21. <http://dx.doi.org/10.1002/stem.1529>
196. Siclari A, Mascaro G, Gentili C, Kaps C, Cancedda R, Boux E. Cartilage repair in the knee with subchondral drilling augmented with a platelet-rich plasma-immersed polymer-based implant. *Knee Surg Sports Traumatol Arthrosc* 2014;**22**:1225–34. <http://dx.doi.org/10.1007/s00167-013-2484-1>
197. Arthrosurface®. *Arthrosurface Knee Products*. 2014. URL: www.arthrosurface.com/products/knee/ (accessed 20 July 2016).
198. Arthrosurface®. *Patello Femoral Implants*. 2014. URL: www.arthrosurface.com/products/knee/patello-femoral-implants/ (accessed 25 July 2016).
199. Arthrosurface®. *Unicondylar Implants*. 2014. URL: www.arthrosurface.com/products/knee/unicondylar-implants/ (accessed 25 July 2016).
200. Arthrosurface®. *Femoral Condyle Implants*. 2014. URL: www.arthrosurface.com/products/knee/femoral-condyle-implants-intl-only/ (accessed 25 July 2016).
201. BioPoly™. *BioPoly RS Knee System*. 2014. URL: www.biopolyortho.com/Products.aspx (accessed 25 July 2016).
202. Episurf Medical. *The Episealer Trochlea*. 2014. URL: <http://episurf.com/products-3/> (accessed 25 July 2016).
203. Schmitt LC, Quatman CE, Paterno MV, Best TM, Flanigan DC. Functional outcomes after surgical management of articular cartilage lesions in the knee: a systematic literature review to guide postoperative rehabilitation. *J Orthop Sports Phys Ther* 2014;**44**:565–A10. <http://dx.doi.org/10.2519/jospt.2014.4844>
204. Martinčič D, Radosavljevič D, Drobnič M. Ten-year clinical and radiographic outcomes after autologous chondrocyte implantation of femoral condyles. *Knee Surg Sports Traumatol Arthrosc* 2014;**22**:1277–83. <http://dx.doi.org/10.1007/s00167-013-2778-3>
205. Niethammer TR, Müller PE, Safi E, Ficklscherer A, Roßbach BP, Jansson V, Pietschmann MF. Early resumption of physical activities leads to inferior clinical outcomes after matrix-based autologous chondrocyte implantation in the knee. *Knee Surg Sports Traumatol Arthrosc* 2014;**22**:1345–52. <http://dx.doi.org/10.1007/s00167-013-2583-z>
206. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;**273**:408–12. <http://dx.doi.org/10.1001/jama.1995.03520290060030>
207. Deeks JJ, Dines J, D’Amico R, Sowden AJ, Sakarovitch C, Song F, et al. Evaluating non-randomised intervention studies. *Health Technol Assess* 2003;**7**(27). <http://dx.doi.org/10.3310/hta7270>

208. Brix M, Chiari C, Nehrer S, Windhager R, Domayer S. Five to ten year follow up of matrix associated autologous chondrocyte transplantation (MACT) in the knee. *Osteoarthr Cartil* 2012;**20**(Suppl. 1):27. <https://doi.org/10.1016/j.joca.2012.02.547>
209. Briggs KK, Rodkey WG, Steadman J. The effect of lesion size on clinical outcomes after microfracture: a 2 to 8 year follow-up study. *Osteoarthr Cartil* 2013;**21**:S117. <https://doi.org/10.1016/j.joca.2013.02.249>
210. Ebert JR, Smith A, Wood DJ, Ackland TR. A comparison of the responsiveness of 4 commonly used patient-reported outcome instruments at 5 years after matrix-induced autologous chondrocyte implantation. *Am J Sports Med* 2013;**41**:2791–9. <http://dx.doi.org/10.1177/0363546513502314>
211. Ebert JR, Fallon M, Robertson W, Ackland T, Zheng MH, Janes G, *et al.* Five year clinical and radiological evaluation of matrix-induced autologous chondrocyte implantation (MACI). *Arthroscopy* 2011;**1**:e170. <https://doi.org/10.1016/j.arthro.2011.08.152>
212. Ebert JR, Smith A, Edwards PK, Hambly K, Wood DJ, Ackland TR. Factors predictive of outcome 5 years after matrix-induced autologous chondrocyte implantation in the tibiofemoral joint. *Am J Sports Med* 2013;**41**:1245–54. <http://dx.doi.org/10.1177/0363546513484696>
213. Filardo G, Kon E, Di Martino A, Patella S, Altadonna G, Balboni F, *et al.* Second-generation arthroscopic autologous chondrocyte implantation for the treatment of degenerative cartilage lesions. *Knee Surg Sports Traumatol Arthrosc* 2012;**20**:1704–13. <http://dx.doi.org/10.1007/s00167-011-1732-5>
214. Filardo G, Kon E, Andriolo L, Di Matteo B, Balboni F, Marcacci M. Clinical profiling in cartilage regeneration: prognostic factors for midterm results of matrix-assisted autologous chondrocyte transplantation. *Am J Sports Med* 2014;**42**:898–905. <http://dx.doi.org/10.1177/0363546513518552>
215. Filardo G, Kon E, Andriolo L, Di Martino A, Zaffagnini S, Marcacci M. Treatment of ‘patellofemoral’ cartilage lesions with matrix-assisted autologous chondrocyte transplantation: a comparison of patellar and trochlear lesions. *Am J Sports Med* 2014;**42**:626–34. <http://dx.doi.org/10.1177/0363546513510884>
216. Health Quality Ontario. Arthroscopic lavage and debridement for osteoarthritis of the knee: an evidence-based analysis. *Ontario Health Technology Assessment Series* 2005;**5**:1–37.
217. Kon E, Patella S, Filardo G, Di Martino A, Perdisa F, Di Matteo B, *et al.* Autologous chondrocyte transplantation: perspective study at 7-year minimum follow-up. *J Orthop Traumatol* 2011;**12**:S76.
218. Kon E, Filardo G, Condello V, Collarile M, Di Martino A, Zorzi C, Marcacci M. Second-generation autologous chondrocyte implantation: results in patients older than 40 years. *Am J Sports Med* 2011;**39**:1668–75. <http://dx.doi.org/10.1177/0363546511404675>
219. Kon E, Verdonk P, Condello V, Delcogliano M, Dhollander A, Filardo G, *et al.* Matrix-assisted autologous chondrocyte transplantation for the repair of cartilage defects of the knee: systematic clinical data review and study quality analysis. *Am J Sports Med* 2009;**37**(Suppl. 1):156–66. <http://dx.doi.org/10.1177/0363546509351649>
220. Kreuz PC, Steinwachs MR, Ergelet C, Krause SJ, Konrad G, Uhl M, Südkamp N. Results after microfracture of full-thickness chondral defects in different compartments in the knee. *Osteoarthr Cartil* 2006;**14**:1119–25. <http://dx.doi.org/10.1016/j.joca.2006.05.003>
221. McNickle AG, L'Heureux DR, Yanke AB, Cole BJ. Outcomes of autologous chondrocyte implantation in a diverse patient population. *Am J Sports Med* 2009;**37**:1344–50. <http://dx.doi.org/10.1177/0363546509332258>
222. Minas T, Peterson L. Autologous chondrocyte transplantation. *Oper Tech Sports Med* 2012;**20**:72–86.

223. Nawaz S, Dhinsa B, Gallagher K, Carrington R, Skinner J, Briggs T, Bentley G. Autologous chondrocyte implantation does not prevent the need for arthroplasty in patients with pre-existing osteoarthritis. *Arthroscopy* 2011;**27** (Suppl.):e170–1. <https://doi.org/10.1016/j.arthro.2011.08.153>
224. Niemeyer P, Lenz P, Jreuz PC, Salzmann GM, Sudkamp NP, Schal H, *et al.* Chondrocyte-seeded type I/III collagen membrane for autologous chondrocyte transplantation: prospective 2-year results in patients with cartilage defects of the knee joint. *Arthroscopy* 2010;**26**:1074–82. <https://doi.org/10.1016/j.arthro.2009.12.028>
225. Noyes FR, Barber-Westin SD. Advanced patellofemoral cartilage lesions in patients younger than 50 years of age: is there an ideal operative option? *Arthroscopy* 2013;**29**:1423–36. <http://dx.doi.org/10.1016/j.arthro.2013.03.077>
226. Oussedik S, Tsitskaris K, Parker D. Treatment of articular cartilage lesions of the knee by microfracture or autologous chondrocyte implantation: a systematic review. *Arthroscopy* 2015;**31**:745. <https://doi.org/10.1016/j.arthro.2014.11.023>
227. Rosenberger RE, Gomoll AH, Bryant T, Minas T. Repair of large chondral defects of the knee with autologous chondrocyte implantation in patients 45 years or older. *Am J Sports Med* 2008;**36**:2336–44. <http://dx.doi.org/10.1177/0363546508322888>
228. Sciarretta FV. 5 to 8 years follow-up of knee chondral defects treated by PVA-H hydrogel implants. *Eur Rev Med Pharmacol Sci* 2013;**17**:3031–8.
229. Scillia AJ, Aune KT, Andrachuk JS, Cain EL, Dugas JR, Fleisig GS, Andrews JR. Return to play after chondroplasty of the knee in National Football League athletes. *Am J Sports Med* 2015;**43**:663–8. <http://dx.doi.org/10.1177/0363546514562752>
230. Ulstein S, Årøen A, Røtterud JH, Løken S, Engebretsen L, Heir S. Microfracture technique versus osteochondral autologous transplantation mosaicplasty in patients with articular chondral lesions of the knee: a prospective randomized trial with long-term follow-up. *Knee Surg Sports Traumatol Arthrosc* 2014;**22**:1207–15. <http://dx.doi.org/10.1007/s00167-014-2843-6>
231. Upmeier H, Brüggengjürgen B, Weiler A, Flamme C, Laprell H, Willich SN. Follow-up costs up to 5 years after conventional treatments in patients with cartilage lesions of the knee. *Knee Surg Sports Traumatol Arthrosc* 2007;**15**:249–57. <https://doi.org/10.1007/s00167-006-0182-y>
232. Wylie JD, Hartley MK, Kapron AL, Aoki SK, Maak TG. What is the effect of matrices on cartilage repair? A systematic review. *Clin Orthop Relat Res* 2015;**473**:1673–82. <http://dx.doi.org/10.1007/s11999-015-4141-0>
233. Zak L, Aldrian S, Wondrasch B, Albrecht C, Marlovits S. Ability to return to sports 5 years after matrix-associated autologous chondrocyte transplantation in an average population of active patients. *Am J Sports Med* 2012;**40**:2815–21. <https://doi.org/10.1177/0363546512462382>

Appendix 1 Search strategies for systematic review and primary studies

Searches for systematic reviews and assessment reports

Cochrane Database of Systematic Reviews: Issue 6 of 12, June 2014

(autologous chondrocyte* near/3 (implant* or transplant*))

Ovid MEDLINE® 1946 to 17 June 2014

1. exp Chondrocytes/tr [Transplantation]
2. exp Cartilage, Articular/tr [Transplantation]
3. exp Transplantation, Autologous/
4. (MACI or MACT or chondroelect or ACI).tw.
5. (chondrocyte* adj4 (implant* or transplant* or matrix or characteri*)).tw.
6. (autologous adj3 (implant* or transplant* or chondrocyte* or cartilage)).tw.
7. (cartilage* adj2 (transplant* or implant*)).tw.
8. 1 or 2 or 3 or 4 or 5 or 6 or 7
9. (systematic review or meta-analysis).tw.
10. meta-analysis.pt.
11. "cochrane database of systematic reviews".jn.
12. 9 or 10 or 11
13. 8 and 12
14. limit 13 to yr="2004 -Current"
15. knee*.af.
16. 14 and 15

EMBASE 1980 to 17 June 2014

1. exp *chondrocyte implantation/
2. (MACI or MACT or chondroelect or ACI).tw.
3. (chondrocyte* adj4 (implant* or transplant* or matrix or characteri*)).tw.
4. (autologous adj3 (implant* or transplant* or chondrocyte* or cartilage)).tw.
5. (cartilage* adj2 (transplant* or implant*)).tw.
6. 1 or 2 or 3 or 4 or 5
7. knee*.af.
8. 6 and 7
9. limit 8 to yr="2004 -Current"
10. systematic review or meta-analysis).tw.
11. 9 and 10

Health Technology Assessment and other assessment reports

Searched the website of the CRD HTA database at www.crd.york.ac.uk/CRDWeb/HomePage.asp and EMA, the US FDA.

Searches for primary studies for clinical effectiveness

Cochrane Central Register of Controlled Trials, Issue 6 of 12, June 2014

(autologous chondrocyte* near/3 (implant* or transplant*))

Ovid MEDLINE 1946 to 17 June 2014

1. exp Chondrocytes/tr [Transplantation]
2. exp Cartilage, Articular/tr [Transplantation]
3. exp Transplantation, Autologous/
4. MACI or MACT or chondrocelect or ACI).tw.
5. (chondrocyte* adj4 (implant* or transplant* or matrix or characteri*)).tw.
6. (autologous adj3 (implant* or transplant* or chondrocyte* or cartilage)).tw.
7. (cartilage* adj2 (transplant* or implant*)).tw.
8. 1 or 2 or 3 or 4 or 5 or 6 or 7
9. Knee/ or knee*.mp.
10. 8 and 9
11. limit 10 to yr="2010 -Current"
12. Animals/
13. Humans/
14. 12 not 13
15. 11 not 14

EMBASE 1947 to 17 June 2014

1. exp *chondrocyte implantation/
2. MACI or MACT or chondrocelect or ACI).tw.
3. (chondrocyte* adj4 (implant* or transplant* or matrix or characteri*)).tw.
4. (autologous adj3 (implant* or transplant* or chondrocyte* or cartilage)).tw.
5. (cartilage* adj2 (transplant* or implant*)).tw.
6. 1 or 2 or 3 or 4 or 5
7. exp knee/
8. knee*.tw.
9. 7 or 8
10. 6 and 9
11. limit 10 to yr="2010-Current"
12. (rat or rats or pig or pigs or porcine or mice or murine or mouse or sheep or rabbit* or canine or dog*).ti.
13. 11 not 12

Web of Science Core Collection 2010 to June 2014

TITLE: (("autologous chondrocyte" or "autologous cartilage") ... More TITLE: (("autologous chondrocyte" or "autologous cartilage") and (implant* or transplant*)) AND TOPIC: (knee*)

TITLE: (MACI or MACT or ACI or condrocelect or "characteri* chondrocyte*") AND TOPIC: (knee*)

Additional searches for other literature**Societies with meetings abstracts available online**

- ISAKOS International Society of Arthroscopy Knee Surgery & Orthopaedic Sports Medicine Biennial Congress 2013: www.isakos.com/
- AAOS American Academy of Orthopaedic Surgeons Annual Meeting: www.aaos.org/Annual
- ORS Orthopaedic Research Society from 1999 to 2014: www.ors.org/abstract-search/
- AOSSM American Orthopaedic Society for Sports Medicine 2013 Annual Meeting: www.sportsmed.org/
- BASK British Association for Surgery of the Knee 2013 abstracts: <http://professional.baskonline.com/content/BASKCurrent.aspx>

Searches for guidelines

- NHS Evidence: www.evidence.nhs.uk/
- British Orthopaedic Association: www.boa.ac.uk/
- BASK: www.baskonline.com/

Ongoing or recently completed studies searched on 3 October 2014

1. ClinicalTrials.gov: <http://clinicaltrials.gov/>
2. World Health Organization (WHO) Clinical Trials Registry Platform Search Portal: <http://apps.who.int/trialsearch/Default.aspx>
3. Current Controlled Trials: www.controlled-trials.com/
4. UK Clinical Trials Gateway: www.ukctg.nihr.ac.uk/default.aspx
5. EU Clinical Trials Register website: www.clinicaltrialsregister.eu/
6. UK Clinical Research Network Study Portfolio: <http://public.ukcrn.org.uk/search/>
7. European Clinical Trials Database (EudraCT): <https://eudract.ema.europa.eu/>

Additional searches

In addition, the inclusion lists of recent systematic reviews were checked.

Auto-alerts in MEDLINE and EMBASE were run for the duration of the review to ensure that newly published studies were identified.

Search strategy for long-term studies

The search strategy below was run in Ovid MEDLINE 1846 to May week 2 2005 and Ovid MEDLINE® In-Process & Other Non-Indexed Citations 15 May.

1. *Cartilage Diseases/su [Surgery]
2. *Arthroplasty, Subchondral/
3. *Cartilage, Articular/su [Surgery]
4. Chondrocytes/tr [Transplantation]
5. microfracture.tw.
6. (chondrocyte* adj4 (implant* or transplant* or matrix or characteri*)).tw.
7. (autologous adj3 (implant* or transplant* or chondrocyte* or cartilage)).tw.
8. (cartilage* adj2 (transplant* or implant*)).tw.
9. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8
10. knee.tw.
11. *Knee Injuries/su [Surgery]
12. 10 or 11
13. 9 and 12
14. limit 13 to yr="1997 -Current"
15. limit 14 to english language

The strategy below was run in Ovid EMBASE 1974 to 15 May 2015:

1. exp microfracture/
2. exp chondrocyte implantation/
3. *Cartilage Diseases/su [Surgery]
4. *Arthroplasty, Subchondral/
5. *Cartilage, Articular/su [Surgery]
6. microfracture.tw.

7. (chondrocyte* adj4 (implant* or transplant* or matrix or characteri*)).tw.
8. (autologous adj3 (implant* or transplant* or chondrocyte* or cartilage)).tw.
9. (cartilage* adj2 (transplant* or implant*)).tw.
10. knee.tw.
11. *Knee Injuries/su [Surgery]
12. 10 or 11
13. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9
14. 12 and 13
15. limit 14 to (english language and yr="1997 -Current")

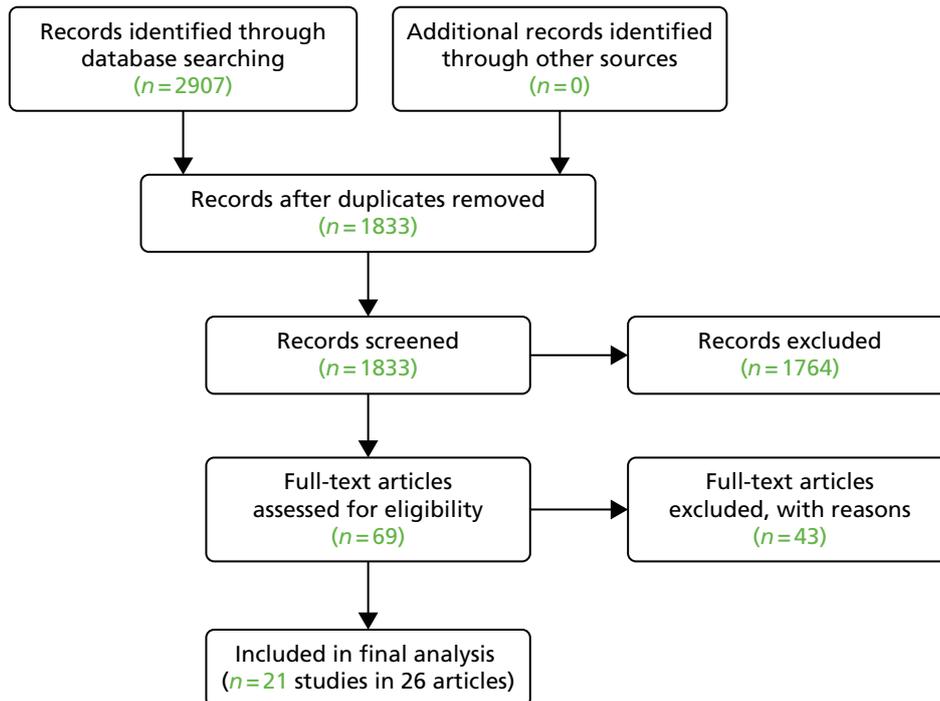


FIGURE 47 Flow diagram for long-term studies.

Appendix 2 Flow diagram systematic review

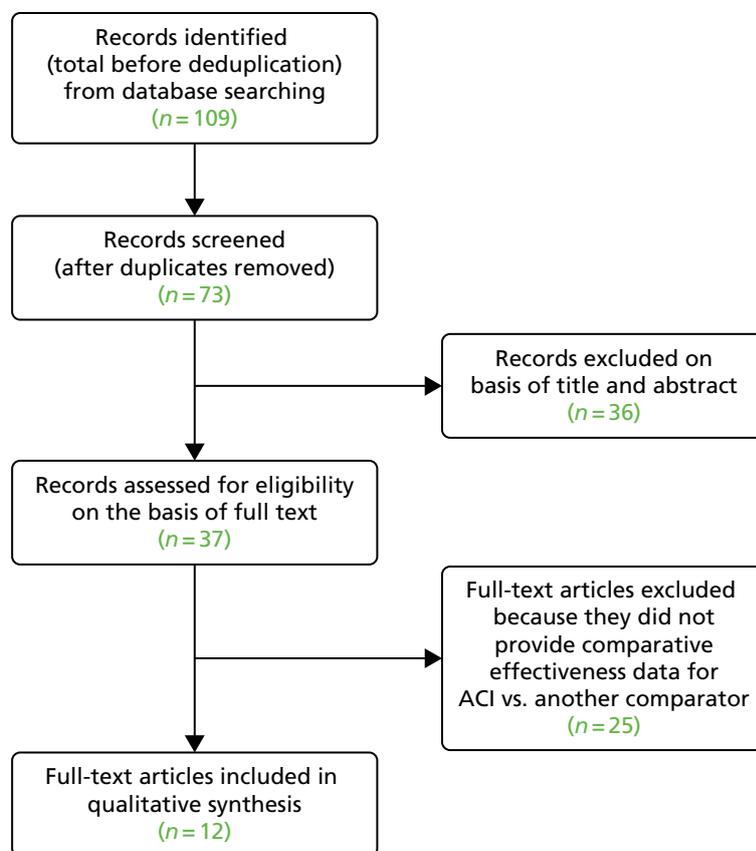


FIGURE 48 A PRISMA study flow diagram for searches for systematic reviews.

Appendix 3 Quality assessment of reviews

Methodology and quality

Most reviews were rated as being of at least medium quality, with three reviews being rated as of low quality (Goyal 2013⁴⁷ and Goyal 2013,⁴⁸ Naveen 2012⁴⁹), six reviews rated as medium quality (Bekkers 2009,⁵⁰ Kon 2009,⁵¹ Magnussen 2008,⁵² Mithöfer 2009,⁵³ Nakamura 2009,⁵⁴ Negrin 2013⁵⁵) and three reviews rated as high quality (Harris 2010,⁴⁴ Vasiliadis 2010,⁴⁵ Vavken 2010⁵⁶).

Ten of the 12 studies had an adequate description of inclusion criteria. Goyal 2013⁴⁷ and Goyal 2013⁴⁸ had no adequate description of participants and outcome measures. Only Harris *et al.* 2010⁴⁴ was rated as having a fully adequate search strategy. Search limitations included 'only PubMed/MEDLINE used' (Goyal 2013⁴⁷ and Goyal 2013,⁴⁸ Nakamura 2009⁵⁴); 'English studies only included' (Goyal 2013⁴⁷ and Goyal 2013,⁴⁸ Kon 2009,⁵¹ Magnussen 2008,⁵² Mithöfer 2009,⁵³ Nakamura 2009⁵⁴); 'limited search terms (limited description or only few terms used)' (Goyal 2013⁴⁷ and Goyal 2013,⁴⁸ Nakamura 2009,⁵⁴ Naveen 2012,⁴⁹ Negrin 2013,⁵⁵ Vavken 2010⁵⁶); and 'no additional searches mentioned' (Bekkers 2009,⁵⁰ Goyal 2013⁴⁷ and Goyal 2013,⁴⁸ Nakamura 2009,⁵⁴ Naveen 2012,⁴⁹ Negrin 2013⁵⁵).

Study selection was adequately described by only four reviews (Harris 2010,⁴⁴ Nakamura 2009,⁵⁴ Negrin 2013,⁵⁵ Vasiliadis 2010⁴⁵); where described, selection was done by independent reviewers. Study flow was adequately shown (or described) by seven reviews (Goyal 2013⁴⁷ and Goyal 2013,⁴⁸ Harris 2010,⁴⁴ Magnussen 2008,⁵² Naveen 2012,⁴⁹ Negrin 2013,⁵⁵ Vavken 2010⁵⁶). Quality assessment was adequately described by eight reviews (Bekkers 2009,⁵⁰ Harris 2010,⁴⁴ Kon 2009,⁵¹ Mithöfer 2009,⁵³ Nakamura 2009,⁵⁴ Negrin 2013,⁵⁵ Vasiliadis 2010,⁴⁵ Vavken 2010⁵⁶); quality assessment tools included the Cochrane risk of bias tool (Bekkers 2009,⁵⁰ Negrin 2013,⁵⁵ Vasiliadis 2010⁴⁵), the Coleman methodology score (modified in some cases) (Bekkers 2009,⁵⁰ Harris 2010,⁴⁴ Kon 2009,⁵¹ Magnussen 2008,⁵² Mithöfer 2009⁵³), the Delphi list (Harris 2010⁴⁴), the rating system of the *Journal of Bone and Joint Surgery* plus Cochrane criteria (Nakamura 2009⁵⁴), a quality scale for observational studies by Deeks (Negrin 2013⁵⁵) and an unnamed list of quality items (Vavken 2010⁵⁶). One review used quality as a basis for further selection (Bekkers 2009⁵⁰). Items for data extraction were listed by eight reviews (Bekkers 2009,⁵⁰ Harris 2010,⁴⁴ Kon 2009,⁵¹ Magnussen 2008,⁵² Mithöfer 2009,⁵³ Naveen 2012,⁴⁹ Vasiliadis 2010,⁴⁵ Vavken 2010⁵⁶) and data extraction was done in duplicate by independent reviewers in five reviews (Kon 2009,⁵¹ Nakamura 2009,⁵⁴ Negrin 2013,⁵⁵ Vasiliadis 2010,⁴⁵ Vavken 2010⁵⁶). Most reviews did not include a meta-analysis and data were summarised in text and tables. A meta-analysis was included in the review by Negrin 2013⁵⁵ and the Cochrane review by Vasiliadis *et al.* 2010.⁴⁵ Some reviews looked for patient characteristics that were related to treatment outcome.

All of the studies described the characteristics of included studies at least to some extent – but a number of reviews did not give details of the quality of individual studies (Goyal 2013⁴⁷ and Goyal 2013,⁴⁸ Kon 2009,⁵¹ Negrin 2013⁵⁵). All reviews showed the results of individual studies – but this was sometimes limited and numerical data were not always reported.

Appendix 4 Studies included in reviews

*T*able 52 shows publications belonging together and referring to the same study population.

TABLE 52 Primary prospective comparative studies in reviews and from extra searches

| Comparison | Bekkers 2009 ⁵⁰ | Goyal 2013 ⁴⁷ | Goyal 2013 ⁴⁸ | Harris 2010 ⁴⁴ | Kon 2009 ⁵¹ | Magnussen 2008 ⁵² | Mithöfer 2009 ⁵³ | Nakamura 2009 ⁵⁴ | Naveen 2012 ⁴⁹ | Negrin 2013 ⁵⁵ | Vasiliadis 2010 ⁵ | Vavken 2010 ⁵⁶ |
|---|-------------------------------|-----------------------------|-----------------------------|------------------------------|---------------------------|---------------------------------|--------------------------------|--------------------------------|------------------------------|------------------------------|---------------------------------|------------------------------|
| C-ACI vs. P-ACI <i>RCT</i> | | | | | | | | | | | | |
| Gooding 2006 ²³ | ✓ | | ✓ | | | | | ✓ | | | ✓ | |
| ACI vs. MACI <i>RCTs</i> | | | | | | | | | | | | |
| Bartlett 2005 ⁵⁷ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | ✓ | | | ✓ | |
| Zeifang 2010 ⁵⁸ | ✓ | ✓ | ✓ | ✓ | | | | | ✓ | | | |
| <i>Comparative cohort</i> | | | | | | | | | | | | |
| Ergelet 2009 ⁵⁹ | | | | | | | | | ✓ | | | |
| Niemeyer 2008 ⁶⁰ | | | | | | | | | ✓ | | | |
| Open vs. arthroscopic ACI <i>Comparative cohort/CCT</i> | | | | | | | | | | | | |
| Ferruzzi 2008 ⁶¹ | | | | ✓ | ✓ | | | | | | | |
| ACI vs. mosaicplasty <i>RCT</i> | | | | | | | | | | | | |
| Bentley 2003 ⁵ | ✓ | | | | | ✓ | | ✓ | ✓ | | ✓ | ✓ |
| Dozin 2005 ⁶² | | | | ✓ | | | | ✓ | ✓ | | ✓ | ✓ |
| <i>CCT</i> | | | | | | | | | | | | |
| ^a Horas 2000, ⁶³ 2003 ⁶ | | | | ✓ | | ✓ | | ✓ | ✓ | | ✓ | ✓ |
| ACI vs. MF <i>RCT</i> | | | | | | | | | | | | |
| Basad 2004, ⁸ 2010; ⁶⁴ Bachmann 2004 ⁶⁵ | | | ✓ | ✓ | ✓ | | | | ✓ | | ✓ | ✓ |
| Crawford 2012 ⁶⁶ | | | | | | | | | | ✓ | | |
| Knutsen 2004, ⁷ 2007 ⁶⁷ | ✓ | | ✓ | ✓ | | ✓ | | ✓ | ✓ | ✓ | ✓ | ✓ |

| Comparison | Bekkers 2009 ⁵⁰ | Goyal 2013 ⁴⁷ | Goyal 2013 ⁴⁸ | Harris 2010 ⁴⁴ | Kon 2009 ⁵¹ | Magnussen 2008 ⁵² | Mithöfer 2009 ⁵³ | Nakamura 2009 ⁵⁴ | Naveen 2012 ⁴⁹ | Negrin 2013 ⁵⁵ | Vasiliadis 2010 ⁴⁵ | Vavken 2010 ⁵⁶ |
|---|-------------------------------|-----------------------------|-----------------------------|------------------------------|---------------------------|---------------------------------|--------------------------------|--------------------------------|------------------------------|------------------------------|----------------------------------|------------------------------|
| Lim 2012⁶⁸ | | ✓ | | | | | | | | | | |
| Saris 2008,⁶⁹ 2009,⁷⁰ Vanlauwe 2011,⁴³ Van Assche 2009,⁷¹ 2010⁷² | ✓ | ✓ | ✓ | ✓ | | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| <i>Comparative cohort</i> | | | | | | | | | | | | |
| Kon 2009⁷³ | | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | |
| Kon 2011⁷⁴ | | ✓ | | | | | | | | | | |
| Minas 2009⁷⁵ | | | | | | | | | ✓ | | | |
| ACI vs. BMSC | | | | | | | | | | | | |
| <i>Comparative cohort</i> | | | | | | | | | | | | |
| Nejadnik 2010⁷⁶ | | | | | | | | | ✓ | | | |
| ACI vs. abrasionplasty | | | | | | | | | | | | |
| <i>RCT</i> | | | | | | | | | | | | |
| Visna 2004⁷⁷ | | | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| BMSC, bone marrow-derived mesenchymal stem cell; C-ACI, collagen-based ACI; MACI, matrix-assisted autologous chondrocyte implantation; P-ACI, periosteum-based ACI. a Described as RCT, but inadequate randomisation method (alternation). | | | | | | | | | | | | |

Appendix 5 Characteristics of systematic reviews

TABLE 53 Characteristics and quality of systematic reviews

| Review | Inclusion criteria and methodology | Included studies | Quality |
|---|--|---|--|
| <p>Bekkers 2009⁵⁰</p> <p><i>Focus:</i> to identify parameters for valid treatment selection in the repair of articular cartilage lesions of the knee</p> <p><i>Funding:</i> not reported, but stated that the authors have no conflicts of interest</p> | <p>INCLUSION CRITERIA</p> <p>Study design: prospective randomised and quasi-randomised trials</p> <p>Participants: focal cartilage lesions of the knee</p> <p>Intervention: comparison of at least two of ACI, MF or osteochondral autologous transplantation</p> <p>Outcomes: not specified</p> <p>METHODOLOGY</p> <p>Search strategy:</p> <p>Databases: PubMed, EMBASE, The Cochrane Library, CENTRAL</p> <p><i>Date of search:</i> 25 August 2009; keywords indicated</p> <p><i>Limitations:</i> PubMed limited by title and abstract, articles in English, German, French or Dutch</p> <p><i>Additional searches:</i> none</p> <p>Study selection: based on titles and abstracts, but not stated how many reviewers were involved</p> <p>Quality assessment: done by two independent reviewers, based on Cochrane risk of bias tool and Coleman methodology score; quality used as a basis for further selection</p> <p>Data extraction: items extracted listed</p> <p>Meta-analysis: no</p> <p>Data analysis: text and tables</p> <p>Subgroup/SAs: looked for indicators of treatment selection/patient profile</p> | <p>No. of included trials: Four (three studies including ACI; only these are considered here)</p> <p>No. of participants: 298</p> <p>TRIALS</p> <p>Design: RCTs</p> <p>Follow-up: 19 months to 5 years</p> <p>Quality: only level of evidence 1b included; one of three studies had some selection, detection and reporting bias; Coleman score 74–94</p> <p>Origin: NR</p> <p>Funding: NR</p> <p>PARTICIPANTS</p> <p>Age: 30.9–33.9 years</p> <p>Sex: 57–80% men</p> <p>Defect size: mean 2.4–5.1 cm²</p> <p>Duration of symptoms: NR</p> <p>Other:</p> <p>Femoral condyles, <i>n</i> = 2</p> <p><i>n</i> = 1; 53% medial femur, 25% patella, 18% lateral femur, 3% trochlea, 1% lateral tibia</p> <p>INTERVENTIONS</p> <ul style="list-style-type: none"> • ACI, <i>n</i> = 2 • CCI, <i>n</i> = 1 • Mosaicplasty, <i>n</i> = 1 • MF, <i>n</i> = 2 <p>OUTCOMES</p> <p>Clinical outcomes (modified Cincinnati, KOOS, Lysholm, VAS, Tegner), SF-36, ICRS macroscopic grading, histology</p> | <p>Inclusion criteria described/adequate: yes</p> <p>Literature search described/adequate: partly, no additional searches</p> <p>Study selection described/adequate: no</p> <p>Data extraction described/adequate: yes</p> <p>Study quality assessment described/adequate: yes</p> <p>Study flow shown: partly, in the text</p> <p>Study characteristics of individual studies described: yes</p> <p>Quality of individual studies given: yes</p> <p>Results of individual studies shown: yes</p> <p>Statistical analysis appropriate: yes</p> <p>OVERALL QUALITY: medium</p> |

continued

TABLE 53 Characteristics and quality of systematic reviews (continued)

| Review | Inclusion criteria and methodology | Included studies | Quality |
|---|--|---|--|
| <p>Goyal 2013⁴⁷</p> <p><i>Focus:</i> to examine the level I and level II evidence for newer generations of ACI vs. first-generation ACI, and to establish if the newer generations have overcome the limitations of first-generation ACI</p> <p><i>Funding:</i> not reported, but stated that the authors have no conflicts of interest</p> | <p>INCLUSION CRITERIA</p> <p>Study design: phase I or II RCTs, systematic reviews/meta-analyses, prospective cohort studies</p> <p>Participants: no criteria specified</p> <p>Intervention: comparison of newer methods of ACI (suspended cultured chondrocytes with covering of collagen membrane; procedures delivering ACI using cell carriers or cell-seeded scaffolds)</p> <p>Outcomes: no criteria specified</p> <p>METHODOLOGY</p> <p>Search strategy:</p> <p>Databases: PubMed</p> <p><i>Date of search:</i> November 2012; keywords listed (partially), including restriction by study type</p> <p><i>Limitations:</i> past 10 years, English language</p> <p><i>Additional searches:</i> not specified</p> <p>Study selection: methods not stated; flow chart shown</p> <p>Quality assessment: no quality assessment reported</p> <p>Data extraction: methods not stated</p> <p>Meta-analysis: no</p> <p>Data analysis: text</p> <p>Subgroup/SAs: none; comparisons described individually</p> | <p>No. of included trials: Seven (four studies comparing interventions; one study comparing younger and older patients; two studies of rehabilitation); only first four studies considered here</p> <p>No. of participants: comparative intervention studies, 180 [only reported for 3 of 4 studies, range 21–91 per study ($n = 3$)]</p> <p>TRIALS</p> <p>Design: comparative intervention studies: 3 RCTs, 1 cost-effectiveness study</p> <p>Follow-up: 1–2 years</p> <p>Quality: not reported; two trials referred to as level I evidence, and two trials as level II evidence</p> <p>Origin: NR</p> <p>Funding: NR</p> <p>PARTICIPANTS</p> <p>Age: mean age 29.3–33.7 years (reported by three RCTs)</p> <p>Defect size: mean 4.1–6 cm² (reported by three RCTs)</p> <p>No further details reported</p> <p>INTERVENTIONS</p> <ul style="list-style-type: none"> • Periosteum-based ACI vs. collagen-based ACI, $n = 2$ • Periosteum-based ACI vs. MACI, $n = 1$ • Collagen-based ACI vs. MACI, $n = 1$ <p>OUTCOMES</p> <p>Clinical and activity scores, cost-effectiveness, QoL, MRI results</p> | <p>Inclusion criteria described/adequate: partially described; inadequate</p> <p>Literature search described/adequate: partially described; inadequate</p> <p>Study selection described/adequate: not described; inadequate</p> <p>Data extraction described/adequate: not described; inadequate</p> <p>Study quality assessment described/adequate: not described; inadequate</p> <p>Study flow shown: yes</p> <p>Study characteristics of individual studies described: yes, but limited</p> <p>Quality of individual studies given: no</p> <p>Results of individual studies shown: yes, but limited</p> <p>Statistical analysis appropriate: n/a</p> <p>OVERALL QUALITY: low</p> |

TABLE 53 Characteristics and quality of systematic reviews (continued)

| Review | Inclusion criteria and methodology | Included studies | Quality |
|---|--|--|---|
| Goyal 2013 ⁴⁸ | INCLUSION CRITERIA | No. of included trials: 15 (11 studies comparing with ACI, only these are considered here – counts separate papers as separate studies, probably just six separate study populations) | Inclusion criteria described/adequate: partially described; inadequate |
| <i>Focus:</i> to examine the level I and level II evidence for MF techniques for cartilage repair | Study design: phase I or II RCTs, systematic reviews/meta-analyses, prospective cohort studies | No. of participants: Six separate ACI study populations: 449 (range=118 per study) | Literature search described/adequate: partially described; inadequate |
| <i>Funding:</i> not reported, but stated that the authors have no conflicts of interest | Participants: no criteria specified | TRIALS | Study selection described/adequate: not described; inadequate |
| | Intervention: MF/marrow stimulation techniques | Design: study types not clearly reported ($n = 4$ RCTs, $n = 2$ comparative cohort) | Data extraction described/adequate: not described; inadequate |
| | Outcomes: no criteria specified | Follow-up: 1.5–7.5 years | Study quality assessment described/adequate: not described; inadequate |
| | METHODOLOGY | Quality: not reported; 5/11 studies referred to as level I evidence and 5/11 as level II evidence | Study flow shown: yes |
| | Search strategy: | Origin: NR | Study characteristics of individual studies described: yes |
| | Databases: PubMed | Funding: NR | Quality of individual studies given: no |
| | <i>Date of search:</i> November 2012; keywords listed (partially), including restriction by study type | PARTICIPANTS | Results of individual studies shown: yes |
| | <i>Limitations:</i> past 10 years, English language | Age: mean age 26.5–37.5 years (one study reported only on range 18–45 years) | Statistical analysis appropriate: n/a |
| | <i>Additional searches:</i> not specified | Sex: unclear, reported only for MF group, more men than women | OVERALL QUALITY: low |
| | Study selection: methods not stated; flow chart shown | Defect size: mean 1.9–2.8 cm ² (two studies reported only ranges 2–10 cm ² and 4–10 cm ²) | |
| | Quality assessment: no quality assessment reported | Duration of symptoms: 1.6–3 years (reported by three studies) | |
| | Data extraction: methods not stated | INTERVENTIONS | |
| | Meta-analysis: no | ACI ($n = 6$) | |
| | Data analysis: text and tables | <ul style="list-style-type: none"> • Characterised chondrocytes, $n = 2$ • MACI, $n = 1$ • Scaffold-based ACI, $n = 2$ • Periosteum-based ACI, $n = 1$ | |
| | Subgroup/SAs: none; comparisons described individually | All vs. MF | |
| | | OUTCOMES | |
| | | Clinical and activity scores, histology | |

continued

TABLE 53 Characteristics and quality of systematic reviews (continued)

| Review | Inclusion criteria and methodology | Included studies | Quality |
|--|--|--|---|
| <p>Harris 2010⁴⁴</p> <p><i>Focus:</i> effect of ACI vs. other cartilage procedures on clinical outcomes, MRI, arthroscopic assessment and durability; effect of different generations of ACI and of patient- and defect-specific parameters</p> <p><i>Funding:</i> no specific funding</p> | <p>INCLUSION CRITERIA</p> <p>Study design: level I and II evidence (RCTs with > 80% FU; RCTs with < 80% FU, prospective cohort studies); minimum duration of FU 12 months</p> <p>Participants: participants with Outerbridge/ICRS grade 3 or 4 focal cartilage defects of the knee</p> <p>Intervention: (1) comparison of any generation ACI with any cartilage repair or restoration technique; (2) comparison of any generation ACI with a different generation of ACI; (3) evaluation of both arthroscopic and open arthroscopy ACI</p> <p>Outcomes: validated clinical outcome measures</p> <p>METHODOLOGY</p> <p>Search strategy:</p> <p>Databases: MEDLINE, EMBASE, CINAHL, PubMed, SPORTDiscus, The Cochrane Library Systematic Reviews</p> <p><i>Date of search:</i> latest search February 2010; keywords listed</p> <p><i>Limitations:</i> no relevant limitations</p> <p><i>Additional searches:</i> bibliographies of reviewed papers</p> <p>Study selection: independent search and selection by four reviewers, agreement by discussion or in case of persistent disagreement by the senior author</p> <p>Quality assessment: yes, Delphi list and modification of the Coleman methodology score</p> | <p>No. of included trials: 13 (but really just 10 distinct trial populations)</p> <p>No. of participants: 917 (700 distinct participants)</p> <p>TRIALS</p> <p>Design:</p> <p>Level I, $n = 6$</p> <p>Level II evidence, $n = 7$ ($n = 7$ RCTs, $n = 3$ CCT/comparative cohort)</p> <p>Follow-up: 1–5 years</p> <p>Quality: mean Coleman methodology score 54/100 (range 36–64); ($n = 7$ fair, $n = 6$ poor)</p> <p>Origin: NR</p> <p>Funding: four studies declared a financial conflict of interest</p> <p>PARTICIPANTS</p> <p>Age: mean age 28.7–34.2 years</p> <p>Sex: NR</p> <p>Defect size: mean 1.9–6.2 cm²</p> <p>Duration of symptoms: 1.75–8.6 years</p> <p>Other: full thickness (100%) and isolated single defects (80–100%); median 88.5% (0–100%) had had previous surgery (reported by 10 studies)</p> <p>INTERVENTIONS</p> <ul style="list-style-type: none"> ● ACI (497 distinct), $n = 604$ ● MF (161 distinct), $n = 271$ ● Osteochondral autograft, $n = 42$ ● ACI: ● Open ACI second-generation (MACI), $n = 4$ ● Open periosteum cover CCI (ChondroCelect), $n = 2$ ● Open periosteum cover ACI, $n = 7$ ● Arthroscopic ACI second generation (Hyalograft C), $n = 2$ ● $n = 2$ open collagen membrane ACI | <p>Inclusion criteria described/adequate: yes</p> <p>Literature search described/adequate: yes</p> <p>Study selection described/adequate: yes</p> <p>Data extraction described/adequate: partly; inadequate</p> <p>Study quality assessment described/adequate: yes</p> <p>Study flow shown: yes</p> <p>Study characteristics of individual studies described: yes</p> <p>Quality of individual studies given: yes</p> <p>Results of individual studies shown: yes</p> <p>Statistical analysis appropriate: yes</p> <p>OVERALL QUALITY: high</p> |

TABLE 53 Characteristics and quality of systematic reviews (continued)

| Review | Inclusion criteria and methodology | Included studies | Quality |
|---|---|---|---|
| | <p>Data extraction: details of extracted outcomes reported, but no details of methodology</p> <p>Meta-analysis: no</p> <p>Data analysis: tables and text; effect sizes calculated</p> <p>Subgroup/SAs: data presented by comparator</p> | <p>OUTCOMES</p> <p>Clinical outcomes (Lysholm, Tegner, KOOS, ICRS, IKDC, modified Cincinnati), SF-36, histology/histomorphology</p> | |
| <p>Kon 2009⁵¹</p> <p><i>Focus:</i> to summarise all studies related to the clinical application of MACI</p> <p><i>Funding:</i> not reported; stated that the authors have no potential conflict of interest</p> | <p>INCLUSION CRITERIA</p> <p>Study design: any</p> <p>Participants: articular cartilage repair of the knee</p> <p>Intervention: second-generation ACI, MACI</p> <p>Outcomes: 'clinical information'</p> <p>METHODOLOGY</p> <p>Search strategy:</p> <p>Databases: MEDLINE, MEDLINE preprints, EMBASE, CINAHL, Life Science Citations, British National Library of Health, CENTRAL</p> <p><i>Date of search:</i> 1 January 1995 to 1 July 2008; keywords indicated</p> <p><i>Limitations:</i> English language</p> <p><i>Additional searches:</i> bibliographies of relevant studies and reviews</p> <p>Study selection: studies selected by three independent reviewers</p> <p>Quality assessment: modified Coleman methodology score</p> <p>Data extraction: data extracted by three independent reviewers; items extracted listed</p> | <p>No. of included trials: 18</p> <p>No. of participants: 731 (range 8–141)</p> <p>TRIALS</p> <p>Design:</p> <ul style="list-style-type: none"> • <i>n</i> = 2 RCTs • <i>n</i> = 3 prospective comparative (but 1 of these is a RCT) • <i>n</i> = 11 prospective cohort studies or case series • <i>n</i> = 2 retrospective case series <p>Follow-up: range 6.5 months to 5 years, median 2 years</p> <p>Quality: mean modified Coleman methodological score (of 100) 53.1 (SD 1.5 (range 33–82))</p> <p>Origin: NR</p> <p>Funding: NR</p> <p>PARTICIPANTS</p> <p>Age: mean age 26.4–37.6 years</p> <p>Sex: NR</p> <p>Defect size: mean 2.4–6.1 cm²</p> <p>Duration of symptoms: NR</p> | <p>Inclusion criteria described/adequate: yes</p> <p>Literature search described/adequate: yes, but English only</p> <p>Study selection described/adequate: yes</p> <p>Data extraction described/adequate: yes</p> <p>Study quality assessment described/adequate: yes</p> <p>Study flow shown: no</p> <p>Study characteristics of individual studies described: yes</p> <p>Quality of individual studies given: no</p> <p>Results of individual studies shown: individual results plotted, but studies not specified</p> <p>Statistical analysis appropriate: unclear, results of comparative studies not reported</p> <p>OVERALL QUALITY: medium</p> |

continued

TABLE 53 Characteristics and quality of systematic reviews (continued)

| Review | Inclusion criteria and methodology | Included studies | Quality |
|------------------------------------|--|---|---|
| | <p>Meta-analysis: no</p> <p>Data analysis: text and tables</p> <p>Subgroup/SAs: none</p> | <p>Other:</p> <ul style="list-style-type: none"> • Traumatic lesions, $n = 334$ cases • Degenerative, $n = 236$ • OCD, $n = 105$ • Other, $n = 56$; 58% on medial femoral condyle, 17% lateral femoral condyle, 12% patella, 7% trochlea, 4% tibial plateau, 2% multiple areas; 63% had had previous surgery, 41% had additional surgery <p>INTERVENTIONS</p> <p>Only reported for ACI:</p> <ul style="list-style-type: none"> • Hyalograft C (Anika Therapeutics, MA, USA), $n = 9^a$ • BioSeed-C (BioTissue Technologies, Freiburg, Germany), $n = 1$ • Atelocollagen (Reprocell Europe, Glasgow, UK), $n = 1$ • BioCart II (Prochon Biotech, Woburn, MA, USA), $n = 1$ • MACI, $n = 4$ • Cartipatch (TBF Tissue Engineering, Bron, France), $n = 1$ • Chondrograft (Primecell, Prague, Czech Republic), $n = 1$ <p>OUTCOMES</p> <p>Clinical outcomes (IKDC subjective, IKDC objective, Lysholm, Cincinnati; Tegner, ICRS subjective and functional, Stanmore, Meyers, VAS scales, KOOS), SF-36, EQ-5D</p> | |
| Magnussen 2008⁵² | <p>INCLUSION CRITERIA</p> <p>Study design: level I and level II studies – prospective comparative studies; minimum 30 participants, minimum FU 1 year</p> <p>Participants: articular cartilage defects of the knee, full-thickness lesions (Outerbridge grade III or IV)</p> | <p>No. of included trials: 6 (5 involving ACI, 1 trial of OAT vs. MF not considered here)</p> <p>No. of participants: 361 (studies involving ACI, range 40–100)</p> <p>TRIALS</p> <p>Design:</p> <ul style="list-style-type: none"> • RCTs, $n = 4$ • CCT, $n = 1$ | <p>Inclusion criteria described/adequate: yes</p> <p>Literature search described/adequate: yes, but English only</p> <p>Study selection described/adequate: no</p> <p>Data extraction described/adequate: partially</p> |

TABLE 53 Characteristics and quality of systematic reviews (continued)

| Review | Inclusion criteria and methodology | Included studies | Quality |
|--|---|--|--|
| <p><i>Funding:</i> authors have received funding from Vanderbilt Sports Medicine research fund, the National Institute of Arthritis and Musculoskeletal Skin Diseases, and the Pfizer Scholars Award in Epidemiology</p> | <p>Intervention: operative treatment with ACI or OAT compared with another method</p> <p>Outcomes: any clinical outcome measures</p> <p>METHODOLOGY</p> <p>Search strategy:</p> <p>Databases: MEDLINE, CENTRAL, EMBASE, CINAHL</p> <p><i>Date of search:</i> 1 January 1966 to 1 January 2007; keywords listed, restricted by study type</p> <p><i>Limitations:</i> English language</p> <p><i>Additional searches:</i> bibliographies of included trials</p> <p>Study selection: methods not stated</p> <p>Quality assessment: modified Coleman methodology score</p> <p>Data extraction: predesigned form used and data extracted listed; no further methodology described</p> <p>Meta-analysis: no</p> <p>Data analysis: text and tables</p> <p>Subgroup/SAs: none</p> | <p>Follow-up: 1–2 years</p> <p>Quality: quality scores for each study not detailed; all included studies were subject to some degree of bias, including selection bias, transfer bias, detection bias</p> <p>Origin: NR</p> <p>Funding: NR</p> <p>PARTICIPANTS</p> <p>Age: mean 30.8–33.5 years</p> <p>Sex: NR</p> <p>Defect size: mean 3.72–6.1 cm²</p> <p>Duration of symptoms: 36–102.7 months (reported by three trials)</p> <p>Other: 43–100% traumatic lesions; 45–89% medial femoral condyle, 10–18% lateral femoral condyle, 0–32% patella, 0–13% trochlea, 0–8% tibial plateau; one trial reported co-interventions; time to full weightbearing 1 day to 12 weeks</p> <p>INTERVENTIONS</p> <p>Every trial examined a different comparison:</p> <ul style="list-style-type: none"> ● C-ACI vs. MACI ● P-ACI vs. MF ● MACI vs. abrasion ● P-ACI or C-ACI vs. open OAT ● P-ACI vs. open OAT <p>OUTCOMES</p> <p>Clinical scoring systems (ICRS, VAS, Stanmore, Lysholm, IKDC, Tegner, Meyers, modified Cincinnati), arthroscopy, histology</p> | <p>Study quality assessment described/adequate: partially</p> <p>Study flow shown: yes, described in the text</p> <p>Study characteristics of individual studies described: yes</p> <p>Quality of individual studies given: not overall, but quality criteria described in the text</p> <p>Results of individual studies shown: yes</p> <p>Statistical analysis appropriate: n/a</p> <p>OVERALL QUALITY: medium</p> |

continued

TABLE 53 Characteristics and quality of systematic reviews (continued)

| Review | Inclusion criteria and methodology | Included studies | Quality |
|---|--|---|---|
| <p>Mithöfer 2009⁵³</p> <p><i>Focus:</i> to assess the effects of articular cartilage repair on athletic participation</p> <p><i>Funding:</i> NR</p> | <p>INCLUSION CRITERIA</p> <p>Study design: RCTs, prospective and retrospective studies with or without a control group with FU data of ≥ 2 years; studies with macroscopic or histological data from second-look arthroscopy > 12 months after surgery; FU $> 80\%$</p> <p>Participants: athletes with articular cartilage lesions (ICRS grade 3 or 4 chondral or osteochondral defects of the knee (femoral condyle, tibia and patellofemoral); ≥ 20 participants</p> <p>Intervention: articular cartilage repair</p> <p>Outcomes: sports activity-related functional outcome scores, ability to return to sports after surgery, ability to continue participation in athletic activity over time</p> <p>METHODOLOGY</p> <p>Search strategy:</p> <p>Databases: MEDLINE, MEDLINE preprints, EMBASE, CINAHL, Life Science Citations, British National Library of Health (including CENTRAL)</p> <p><i>Date of search:</i> 1966 to 31 May 2009; keywords indicated</p> <p><i>Limitations:</i> English language</p> <p><i>Additional searches:</i> bibliographies of relevant studies and reviews; meeting abstracts</p> <p>Study selection: NR</p> <p>Quality assessment: modified Coleman methodology scores</p> <p>Data extraction: items extracted listed, no details of methodology</p> <p>Meta-analysis: no</p> | <p>No. of included trials: 20 (seven including ACI, with six distinct populations – only these are considered here)</p> <p>No. of participants: 535 distinct participants</p> <p>TRIALS</p> <p>Design:</p> <ul style="list-style-type: none"> • RCT, $n = 1$ • Comparative cohort, $n = 1$ • Cohort without comparison group, $n = 2$ • Case series, $n = 2$ <p>Follow-up: 3–5 years</p> <p>Quality: Coleman methodology score 65–100</p> <ul style="list-style-type: none"> • Level 1 evidence, $n = 1$ • Level 2 evidence, $n = 3$ • Level 4 evidence, $n = 2$ <p>Origin: NR</p> <p>Funding: NR</p> <p>PARTICIPANTS</p> <p>Age: overall 29 (SE 6) years; ACI 28 (SE 4) years</p> <p>Sex: NR</p> <p>Defect size: overall, 3.6 (SE 0.4) cm²; ACI, 5.1 (SE 0.8) cm²</p> <p>Duration of symptoms: overall 21 (SE 3 months; ACI 23 (SE 3) months</p> <p>Other: ACI: lesion type: single only 57%, single and multiple 43%; traumatic only 86%, traumatic and degenerative 14; lesion location: femorotibial only 29%, femorotibial and patellofemoral 71%</p> <p>INTERVENTIONS</p> | <p>Inclusion criteria described/adequate: yes</p> <p>Literature search described/adequate: yes, but English only</p> <p>Study selection described/adequate: no</p> <p>Data extraction described/adequate: partly; inadequate</p> <p>Study quality assessment described/adequate: yes</p> <p>Study flow shown: partly in the text</p> <p>Study characteristics of individual studies described: yes</p> <p>Quality of individual studies given: yes</p> <p>Results of individual studies shown: yes</p> <p>Statistical analysis appropriate: yes</p> <p>OVERALL QUALITY: medium</p> |

TABLE 53 Characteristics and quality of systematic reviews (continued)

| Review | Inclusion criteria and methodology | Included studies | Quality |
|--|--|---|---|
| | <p>Data analysis: text, tables, correlations</p> <p>Subgroup/SAs: analysis by comparison</p> | <p>Of 20 studies:</p> <ul style="list-style-type: none"> • ACI, $n = 7$ • MF, $n = 12$ • OAT, $n = 5$ • Allograft, $n = 1$ <p>OUTCOMES</p> <p>Functional outcomes (KOOS, Tegner), return to sports</p> | |
| <p>Nakamura 2009⁵⁴</p> <p><i>Focus:</i> to determine the effectiveness of cell-based therapy for articular cartilage defects of the knee</p> <p><i>Funding:</i> ISAKOS Scientific Committee (presumably)</p> | <p>INCLUSION CRITERIA</p> <p>Study design: RCTs, prospective comparative studies, systematic reviews, case series</p> <p>Participants: symptomatic chondral lesions of the knee</p> <p>Intervention: cell-based therapies</p> <p>Outcomes: no criteria specified</p> <p>METHODOLOGY</p> <p>Search strategy:</p> <p><i>Databases:</i> MEDLINE</p> <p><i>Date of search:</i> 1994 to January 2009; keywords not indicated</p> <p><i>Limitations:</i> English language</p> <p><i>Additional searches:</i> none</p> <p>Study selection: independent selection by three reviewers, differences resolved by discussion</p> <p>Quality assessment: quality assessment according to the rating system of the <i>Journal of Bone and Joint Surgery</i>, supplemented by criteria of The Cochrane Collaboration and Schulz 1995;²⁰⁶ data evaluated by reviewers independently, differences resolved by discussion</p> <p>Data extraction: data evaluated by reviewers independently, differences resolved by discussion</p> | <p>No. of included trials: 12 ($n = 10$ comparing interventions, $n = 2$ regarding activity levels/ rehabilitation), plus three systematic reviews – only 10 studies comparing interventions considered here ($n = 9$ with distinct populations)</p> <p>No. of participants: 754 in intervention studies reported (really 674 distinct participants)</p> <p>TRIALS</p> <p>Design:</p> <ul style="list-style-type: none"> • RCTs, $n = 9$ • CCT, $n = 1$ <p>Follow-up: 1–5 years</p> <p>Quality:</p> <ul style="list-style-type: none"> • RCTs classified as level I evidence, $n = 2$ • RCTs, $n = 6$ • CCT as level II evidence, $n = 1$ <p>Quality limitations included lack of allocation concealment, not enough information on losses to FU and blinding</p> <p>Origin: NR</p> <p>Funding: NR</p> <p>PARTICIPANTS</p> <p>Age: mean 28.7–33.5 years</p> <p>Sex: NR</p> <p>Defect size: mean 1.9–6 cm²</p> | <p>Inclusion criteria described/adequate: yes</p> <p>Literature search described/adequate: partly; inadequate</p> <p>Study selection described/adequate: yes</p> <p>Data extraction described/adequate: partly; inadequate</p> <p>Study quality assessment described/adequate: yes</p> <p>Study flow shown: no</p> <p>Study characteristics of individual studies described: yes</p> <p>Quality of individual studies given: yes</p> <p>Results of individual studies shown: yes</p> <p>Statistical analysis appropriate: n/a</p> <p>OVERALL QUALITY: medium</p> |

continued

TABLE 53 Characteristics and quality of systematic reviews (continued)

| Review | Inclusion criteria and methodology | Included studies | Quality |
|--|--|--|---|
| | <p>Meta-analysis: no</p> <p>Data analysis: text and tables</p> <p>Subgroup/SAs: data presented by comparator</p> | <p>Duration of symptoms: NR</p> <p>Other: 36–100% traumatic lesions; 24–89% medial femoral condyle, 8.5–23% lateral femoral condyle, 0–61% patella, 0–15.2% trochlea, 0–10% lateral tibial condyle/tibial plateau</p> <p>INTERVENTIONS</p> <p>ACI:</p> <ul style="list-style-type: none"> • P-ACI, $n = 5$ • C-ACI, $n = 3$ • Characterised ACI, $n = 1$ • Hyalograft C, $n = 1$ • MACI, $n = 2$ • OAT, $n = 3$ • MF, $n = 3$ • Abrasion, $n = 1$ <p>OUTCOMES</p> <p>Clinical outcomes (modified Cincinnati, Stanmore, ICRS, IKDC, KOOS, Lysholm, Meyers, Tegner, VAS), SF-36; histology</p> | |
| <p>Naveen 2012⁴⁹</p> <p><i>Focus:</i> to determine the effectiveness of ACI when compared with other treatment modalities</p> <p><i>Funding:</i> no specific funding</p> | <p>INCLUSION CRITERIA</p> <p>Study design: RCTs</p> <p>Participants: no criteria specified</p> <p>Intervention: ACI vs. other treatment modalities (MF, mosaicplasty, abrasionplasty, BMSC, MACI) for cartilage repair in the knee</p> <p>Outcomes: clinical outcomes and evaluation scores; histological outcomes</p> <p>METHODOLOGY</p> <p>Search strategy:</p> <p>Databases: PubMed, Scopus, NICE, CCTR</p> <p><i>Date of search:</i> up to June 2010; only two keywords searched</p> <p><i>Limitations:</i> none</p> <p><i>Additional searches:</i> none</p> | <p>No. of included trials: 17 (but only 13 separate trial populations)</p> <p>No. of participants: 1644 (range 21–321 per study) (number as stated by authors, only 1339 distinct participants)</p> <p>TRIALS</p> <p>Design: not specified ($n = 7$ RCTs, $n = 6$ CCT/comparative cohort)</p> <p>Follow-up: 12 months to 5 years</p> <p>Quality:</p> <ul style="list-style-type: none"> • Classified as level I evidence, $n = 4$ • Level II, $n = 8$ • Level III, $n = 1$ • Level IV, $n = 2$ • No classification, $n = 2$ <p>Origin: NR</p> <p>Funding: NR</p> | <p>Inclusion criteria described/adequate: yes</p> <p>Literature search described/adequate: partly; inadequate</p> <p>Study selection described/adequate: no</p> <p>Data extraction described/adequate: no</p> <p>Study quality assessment described/adequate: no</p> <p>Study flow shown: yes</p> <p>Study characteristics of individual studies described: yes</p> <p>Quality of individual studies given: limited</p> <p>Results of individual studies shown: yes</p> <p>Statistical analysis appropriate: n/a</p> <p>OVERALL QUALITY: low</p> |

TABLE 53 Characteristics and quality of systematic reviews (continued)

| Review | Inclusion criteria and methodology | Included studies | Quality |
|----------------------------------|---|---|---|
| | <p>Study selection: as per inclusion criteria, methods not stated (but obviously actual inclusion was different from inclusion criteria)</p> <p>Quality assessment: limited, for histological assessments, reported blinding of assessors, attrition and level of evidence</p> <p>Data extraction: limited; brief note on items extracted but not methodology</p> <p>Meta-analysis: no</p> <p>Data analysis: text and tables</p> <p>Subgroup/SAs: ACI vs. different comparators</p> | <p>PARTICIPANTS</p> <p>Age: NR</p> <p>Sex: 57–76% male (reported by 14 studies)</p> <p>Defect size: mean 1.9–6.4 cm²</p> <p>INTERVENTIONS</p> <p>ACI vs.:</p> <ul style="list-style-type: none"> • Mosaicplasty, <i>n</i> = 4 • MF, <i>n</i> = 8 • MACI, <i>n</i> = 3 • BMSC, <i>n</i> = 1 • Abrasionplasty, <i>n</i> = 1 <p>OUTCOMES</p> <p>Clinical scores (subjective outcome, Lysholm, Tegner, Cincinnati, Stanmore, Meyers, IHC, ICRS, IKDC, Hop test, KOOS, Gillquist), QoL (SF-36), histology/MRI</p> | |
| Negrin 2013 ⁵⁵ | <p>INCLUSION CRITERIA</p> <p>Study design: CCT or controlled prospective observational study, FU ≥ 1 year</p> <p>Participants: patients with full-thickness cartilage defects (Outerbridge grades III and IV) on the medial or lateral femoral condyle, the trochlea, or the patella due to acute or repetitive trauma, osteonecrosis, or OCD</p> <p>Intervention: MF (without implantation of a scaffold or injection of substitutes) vs. any type of ACI</p> <p>Outcomes: clinical scores (functional capacity)</p> <p>METHODOLOGY</p> <p>Search strategy:</p> | <p>No. of included trials: 6</p> <p>No. of participants: 399</p> <p>TRIALS</p> <p>Design:</p> <ul style="list-style-type: none"> • RCTs, <i>n</i> = 4 • Comparative cohort, <i>n</i> = 2 <p>Follow-up: 1–5 (7.5?) years</p> <p>Quality: NR</p> <p>Origin: NR</p> <p>Funding: NR</p> <p>PARTICIPANTS</p> <p>Age: mean 25.1–40.4 years</p> <p>Sex: NR</p> <p>Defect size: mean 2.0–4.8 cm²</p> <p>Duration of symptoms: NR</p> | <p>Inclusion criteria described/adequate: yes</p> <p>Literature search described/adequate: yes; inadequate</p> <p>Study selection described/adequate: yes</p> <p>Data extraction described/adequate: yes</p> <p>Study quality assessment described/adequate: yes</p> <p>Study flow shown: yes</p> <p>Study characteristics of individual studies described: yes</p> <p>Quality of individual studies given: no</p> <p>Results of individual studies shown: yes</p> |

continued

TABLE 53 Characteristics and quality of systematic reviews (continued)

| Review | Inclusion criteria and methodology | Included studies | Quality |
|--|--|---|---|
| | <p>Databases: MEDLINE, EMBASE; CINAHL, CENTRAL</p> <p><i>Date of search:</i> up to 31 March 2013; only one search term</p> <p><i>Limitations:</i> none</p> <p><i>Additional searches:</i> none</p> <p>Study selection: studies selected by two independent reviewers using standardised forms; discrepancies resolved by consensus</p> <p>Quality assessment: for RCTs, Cochrane risk of bias tool; for observational studies criteria proposed by Deeks <i>et al.</i> 2003,²⁰⁷ assessment by two independent reviewers, discrepancies resolved by consensus</p> <p>Data extraction: items extracted are listed</p> <p>Meta-analysis: yes</p> <p>Data analysis: SMD (random-effects model), heterogeneity, funnel plot; text and tables</p> <p>Subgroup/SAs: duration of FU; generation of ACI</p> | <p>No other characteristics systematically reported</p> <p>INTERVENTIONS</p> <ul style="list-style-type: none"> • First-generation ACI, <i>n</i> = 1 • Second-generation ACI, <i>n</i> = 4 • Third-generation ACI, <i>n</i> = 1 <p>All vs. MF</p> <p>OUTCOMES</p> <p>Clinical outcome (Lysholm, IKDC, KOOS), treatment failure, histology</p> | <p>Statistical analysis appropriate: no; meta-analyses show substantial heterogeneity, which was not explored</p> <p>OVERALL QUALITY: medium</p> |
| <p>Vasiliadis 2010⁴⁵</p> <p><i>Focus:</i> to assess the effectiveness and safety of ACI compared with other treatment options (conservative or surgical) for patients who require knee repair of clinically significant symptomatic defects of the knee joint</p> <p><i>Funding:</i> NR</p> <p>Note: refers to a 2010 Cochrane review, which is slightly less inclusive [three of the trials included here were excluded in the Cochrane review (two trials were comparisons</p> | <p>INCLUSION CRITERIA</p> <p>Study design: RCTs or quasi-randomised trials</p> <p>Participants: 15–55 years with symptomatic cartilage defects of the femur or patella (in joints free from rheumatoid arthritis, OA)</p> <p>Intervention: ACI vs. any other intervention</p> <p>Outcomes: clinical efficacy and complications</p> <p>METHODOLOGY</p> <p>Search strategy:</p> <p>Databases: Cochrane Bone, Joint and Muscle Trauma Group specialised register, CENTRAL, MEDLINE, EMBASE,</p> | <p>No. of included trials: 9</p> <p>No. of participants: 626 (19–118 per study)</p> <p>TRIALS</p> <p>Design:</p> <ul style="list-style-type: none"> • RCTs, <i>n</i> = 8 • Quasi-RCT, <i>n</i> = 1 <p>Follow-up: 10 months to 5 years</p> <p>Quality: overall, average to low quality:</p> <ul style="list-style-type: none"> • < 75% adequate sequence generation • < 50% adequate allocation concealment • < 75% incomplete outcome data addressed | <p>Inclusion criteria described/adequate: yes</p> <p>Literature search described/adequate: yes, but no additional searches mentioned</p> <p>Study selection described/adequate: yes</p> <p>Data extraction described/adequate: yes</p> <p>Study quality assessment described/adequate: yes</p> <p>Study flow shown: no</p> <p>Study characteristics of individual studies described: yes</p> |

TABLE 53 Characteristics and quality of systematic reviews (continued)

| Review | Inclusion criteria and methodology | Included studies | Quality |
|---|---|--|--|
| of different forms of ACI, one trial was excluded because of the heterogeneous patient population]) | <p>SPORTDiscus, WHO International Clinical Trials Registry Platform, CCT</p> <p><i>Date of search:</i> December 2009; reference for search strategy given</p> <p><i>Limitations:</i> none</p> <p><i>Additional searches:</i> none</p> <p>Study selection: independently by two reviewers, differences resolved by discussion</p> <p>Quality assessment: Cochrane risk of bias tool, similarity at baseline; quality assessed by two reviewers independently, differences resolved by discussion</p> <p>Data extraction: items extracted reported; authors contacted for missing information; data extracted by two reviewers independently, differences resolved by discussion</p> <p>Meta-analysis: no/limited for Cochrane review</p> <p>Data analysis: text and tables</p> <p>Subgroup/SAs: none</p> | <p>Origin: NR</p> <p>Funding: NR</p> <p>PARTICIPANTS</p> <p>Age: mean 29.7–35.4 years</p> <p>Sex: 47–68% male</p> <p>Defect size: mean 1.9–6.1 cm²</p> <p>Duration of symptoms: 1.5–10 years</p> <p>Other: location (reported by $n = 7$): medial femoral condyle 24–89%, lateral femoral condyle 5–25%, trochlea 0–21%, patella 0–61%, tibial plateau 0–10%, multiple 0–13%; Aetiology ($n = 5$): trauma 36 to 92%, OCD 8–28%, chondromalacia patellae 0–46%, failed previous surgery 0–20%, uncertain 3–31%</p> <p>INTERVENTIONS</p> <ul style="list-style-type: none"> • ACI, $n = 1$ • C-ACI, $n = 3$ • P-ACI, $n = 5$ • CCI, $n = 1$ • MACI, $n = 1$ • MF, $n = 2$ • Mosaicplasty, $n = 3$ • Abrasion, $n = 1$ <p>OUTCOMES</p> <p>Clinical outcomes (Lysholm, Tegner, KOOS, modified Cincinnati, VAS, Mayers, ICRS, Stanmore), SF-36, biopsy, IKDC, complications</p> | <p>Quality of individual studies given: yes</p> <p>Results of individual studies shown: yes</p> <p>Statistical analysis appropriate: yes</p> <p>OVERALL QUALITY: high</p> |
| Vavken 2010 ⁵⁶ | <p>INCLUSION CRITERIA</p> <p>Study design: controlled trials, minimum FU 6 months</p> <p>Participants: cartilage defects of the knee</p> <p>Intervention: ACI (any type) vs. another cartilage repair procedure or placebo</p> <p>Outcomes: clinical outcome, quality of repair tissue</p> | <p>No. of included trials: 10 (but really only seven independent trials)</p> <p>No. of participants: 441 (range 19–118 per study)</p> <p>TRIALS</p> <p>Design:</p> <ul style="list-style-type: none"> • RCTs, $n = 6$ • Quasi-RCT, $n = 1$ | <p>Inclusion criteria described/adequate: yes</p> <p>Literature search described/adequate: yes, although limited search terms</p> <p>Study selection described/adequate: partly</p> <p>Data extraction described/adequate: yes</p> |

continued

TABLE 53 Characteristics and quality of systematic reviews (continued)

| Review | Inclusion criteria and methodology | Included studies | Quality |
|--------|--|---|---|
| | <p>METHODOLOGY</p> <p>Search strategy:</p> <p>Databases: PubMed; EMBASE, CENTRAL, CINAHL, BioMed</p> <p><i>Date of search:</i> December 2009; search strategy shown</p> <p><i>Limitations:</i> none</p> <p><i>Additional searches:</i> bibliographies of relevant papers</p> <p>Study selection: records compared against inclusion criteria, but no further methodology reported</p> <p>Quality assessment: level of evidence determined, quality criteria listed, independent assessment by two reviewers</p> <p>Data extraction: items extracted listed; independent extraction by two reviewers</p> <p>Meta-analysis: no</p> <p>Data analysis: text and tables</p> <p>Subgroup/SAs: results reported by comparator</p> | <p>Follow-up: 1–5 years</p> <p>Quality:</p> <ul style="list-style-type: none"> Level I evidence, $n = 3$ Level II evidence, $n = 4$; attrition 0–28%, deficits with respect to sample size, randomisation procedure, blinding of outcome assessment <p>Origin: NR</p> <p>Funding: NR</p> <p>PARTICIPANTS</p> <p>Age: NR</p> <p>Sex: 57–68% male</p> <p>Defect size: mean 1.9–5.1 cm²</p> <p>Duration of symptoms: NR</p> <p>Other: NR</p> <p>INTERVENTIONS</p> <p>ACI vs.:</p> <ul style="list-style-type: none"> Osteochondral graft transfer, $n = 3$ MF, $n = 3$ Abrasion, $n = 1$ <p>OUTCOMES</p> <p>Clinical outcome (subjective, Lysholm, Tegner, Meyer, modified Cincinnati, Stanmore, IKDC, KOOS), SF-36, histology, safety</p> | <p>Study quality assessment described/adequate: yes</p> <p>Study flow shown: yes</p> <p>Study characteristics of individual studies described: yes</p> <p>Quality of individual studies given: yes</p> <p>Results of individual studies shown: yes</p> <p>Statistical analysis appropriate: n/a</p> <p>OVERALL QUALITY: high</p> |

BMSC, bone marrow-derived mesenchymal stem cell; C-ACI, collagen-based ACI; CCT, Current Controlled Trials; CCTR, Cochrane controlled trials reports; CINAHL, Cumulative Index to Nursing and Allied Health Literature; FU, follow-up; ISAKOS; International Society of Arthroscopy, Knee Surgery and Orthopaedic Sports Medicine; n/a, not available; NR, not reported; P-ACI, periosteum-based ACI; SMD, standardised mean difference; WHO, World Health Organization.

a An application for to the EMA for approval of Hyalograft in Europe was withdrawn. URL www.ema.europa.eu/ema/index.jsp?curl=pages/medicines/human/medicines/002657/wapp/Initial_authorisation/human_wapp_000161.jsp&mid=WC0b01ac058001d128 (accessed 7 December 2016).

Appendix 6 Results and conclusions of systematic reviews

TABLE 54 Results and conclusions of systematic reviews

| Review | Outcome | No. of studies | Result of meta-analysis/review | Comments |
|------------------------------|------------------|--|---|--|
| General ACI vs. other | | | | |
| Mithöfer 2009 ⁵³ | Clinical outcome | 1 RCT, 5 non-RCTs with ACI, 20 studies overall | Good and excellent results in 82 (SE 7)% [vs. 79 (SE 5)% for all methods] | Increase in Tegner activity score was seen in 84 SE6% of patients overall, the highest average Tegner scores were found for ACI; decreasing Tegner scores were seen in six studies after initial increase: five after MF and one after OAT, no decrease seen with ACI (36–60 months) |
| | Return to sports | 1 RCT, 5 non-RCTs with ACI, 20 studies overall | Return to sports, 33–96% with ACI [mean 67 (SE 17)% vs. 73 (SE 7)% for all methods] | Time to return to sports, 18 (SE 4) months after ACI (range 12–36 months) vs. 8 (SE 1) months after MF, 7 (SE 2) months after osteochondral autograft |
| | Subgroups | 1 RCT, 5 non-RCTs with ACI, 20 studies overall | Better results with younger age (< 25–30 years) | Better results with shorter time between diagnosis and surgical treatment (< 12 months) |
| | | | | In patients treated with ACI: lower average number of previous surgeries in those who returned to sports; return to sports significantly better and time to return significantly shorter in competitive than recreational athletes |

continued

TABLE 54 Results and conclusions of systematic reviews (continued)

| Review | Outcome | No. of studies | Result of meta-analysis/review | Comments |
|-------------------------------|-----------------------|-----------------------------|---|--|
| Vasiliadis 2010 ⁴⁵ | Subgroups | 4 RCTs | <p>1 RCT (1 year): no significant difference by anatomical site, but none of the patellar lesions had a good arthroscopic result</p> <p>1 RCT (1 year): patients with previous surgical procedures had worse clinical outcomes, but correlation not statistically significant; longer duration of symptoms before surgery (C-ACI or MACI) significantly correlated to worse clinical outcomes; patients < 35 years had significantly better clinical outcomes</p> <p>1 RCT: onset of symptoms < 2 years before surgery associated with larger improvement in KOOS (MF and CCI, < 3 years in the latter group)</p> <p>1 RCT (2 years): patients < 30 and more active patients had better results; patients with smaller lesions (< 4 cm²) had better results in the MF group only (result independent of lesion size with P-ACI)</p> | |
| C-ACI vs. P-ACI | | | | |
| Goyal 2013 ⁴⁷ | General effectiveness | 1 RCT | No statistical difference in results after 2 years | Outcomes not specified; actual values not reported for any of the outcomes |
| | Retreatment | 1 RCT, 1 cost-effectiveness | Significant number of patients in P-ACI group required periosteal shaving; high risk of patch hypertrophy | |
| | Cost-effectiveness | 1 cost-effectiveness | Both methods cost-effective but C-ACI slightly more so because of risk of hypertrophy with P-ACI | |
| Harris 2010 ⁴⁴ | Clinical outcome | 1 RCT | No significant difference in modified Cincinnati (2 years) and ICRS AKS (1 and 2 years) scores | |
| | Histology | 1 RCT | No significant difference in macroscopic and histological examination at 1 and 2 years, but 36% in the P-ACI group vs. 0% in the C-ACI group needed arthroscopic knee surgery because of hypertrophy at 1 year | |
| Nakamura 2009 ⁵⁴ | Clinical outcome | 1 RCT | No significant difference in modified Cincinnati and ICRS AKS scores (2 years) | |
| | Histology | 1 RCT | Significant number of patients in P-ACI group required shaving of hypertrophied graft | |

TABLE 54 Results and conclusions of systematic reviews (continued)

| Review | Outcome | No. of studies | Result of meta-analysis/review | Comments |
|---|---------------------------------------|---|--|--|
| Vasiladis 2010 ⁵⁶ | Clinical outcome | 1 RCT | No significant difference in modified Cincinnati score at 2 years (good and excellent results in 66.7% with C-ACI and 74.3% with P-ACI) | |
| | Histology | 1 RCT | 81% good to excellent results with P-ACI and 79% with C-ACI according to ICRS evaluation system (1 year; $p = NS$), but biopsies better for C-ACI (statistical significance unclear) | |
| | Complications | 1 RCT | 12/31 (1 year) and 1/9 (2 years) graft hypertrophies with P-ACI, 1/35 (2 years) with C-ACI | |
| General MACI and second-generation ACI | | | | |
| Kon 2009 ⁵¹ | Clinical outcome | 18 studies (incl. 2 RCTs, 3 additional comparative studies) | Mean subjective preoperative IKDC score ranged from 37.0 to 41.1 and improved to 70.2–80.2 at 5 years (results at earlier time points 73.6–80.6) Mean preoperative Lysholm score ranged from 46.3 to 57.5 and improved to 80.8 at 3 years (results at earlier time points 69.7–96.7) | |
| | Complications | 8 studies | <ul style="list-style-type: none"> • $n = 7$ graft hypertrophy (4 for MACI, 2 for Hyalograft C, 1 for BioSeed) • $n = 4$ joint stiffness (3 for MACI and 1 for Hyalograft C) • $n = 1$ graft detachment for MACI • $n = 1$ synovitis for Hyalograft C <p>One study reported $n = 3$ hypertrophy, $n = 3$ graft detachments and $n = 1$ partial ossification with atelocollagen scaffold (only product used in conjunction with a periosteal flap; impossible to determine if the complication was related to the periosteal flap)</p> | |
| ACI vs. MACI | | | | |
| Naveen 2012 ⁴⁹ | Clinical outcome | 1 RCT, 2 comparative cohort | For 1 RCT and 1 comparative cohort, no significant difference in clinical outcomes (2 years); 1 comparative cohort had significantly better clinical outcomes for MACI, higher complication rate with ACI (4.5 years) | Actual values not reported for any of the outcomes |
| C-ACI vs. MACI | | | | |
| Goyal 2013 ⁴⁷ | Knee function/ clinical scores | 1 RCT | Improvements in all clinical scores with both techniques after 1 year | Actual values not reported for any of the outcomes |
| | Arthroscopic/ histological assessment | 1 RCT | No significant difference after 1 year | |

continued

TABLE 54 Results and conclusions of systematic reviews (continued)

| Review | Outcome | No. of studies | Result of meta-analysis/review | Comments |
|----------------------------------|--------------------------------------|----------------------|---|---|
| Magnussen 2008 ⁵² | Clinical outcome | 1 RCT | No significant difference between groups in modified Cincinnati, VAS and Stanmore scores (1 year) | |
| | Arthroscopic/histological assessment | 1 RCT | ICRS CRA (CRA, 12 = normal cartilage); CRA 8–12, no significant difference (C-ACI 79.2%, MACI 66.6%) (1 year) | |
| | Subgroups | 1 RCT | Per cent with hyaline-like or mixed hyaline/fibrocartilage, no significant difference (C-ACI 42.9%, MACI 36.4%) (1 year) | |
| | Complications | 1 RCT | Patients < 35 years had better clinical outcome ($p = 0.03$) C-ACI: 6.8% arthrofibrosis, 9.1% tissue hypertrophy MACI: 6.4% arthrofibrosis, 6.4% tissue hypertrophy, 2.1% superficial wound infection | |
| Nakamura 2009 ⁵⁴ | Clinical outcome | 1 RCT | No significant difference between groups in modified Cincinnati, VAS, ICRS AKS and Stanmore scores (2 years) | |
| Vasiliadis 2010 ⁴⁵ | Clinical outcome | 1 RCT | No significant difference in modified Cincinnati score (outcome good or excellent in 59.1% after C-ACI, in 72.3% after MACI, 12 months); no significant difference in VAS or Stanmore score | |
| | Histology | 1 RCT | 79.2% good to excellent results with C-ACI and 66.6% with MACI according to ICRS evaluation system (1 year; $p = NS$), with hyaline- or mixed hyaline-like repair tissue in 42.9% with C-ACI and 36.4% with MACI | |
| P-ACI vs. MACI | | | | |
| Goyal 2013 ⁴⁷ | Knee function | 1 RCT | At 2 years, no significant difference between groups in IKDC scores and Tegner activity scores between groups; Lysholm and Gillquist scores (function) favoured P-ACI group | Actual values not reported for most of the outcomes |
| | QoL | 1 RCT | At 2 years, no significant difference between groups in SF-36 scores | |
| | MRI cartilage repair tissue score | 1 RCT | At 1 and 2 years, no significant difference | |
| Harris 2010 ⁴⁴ | Clinical outcome | 2 RCTs | No significant difference in clinical scores after 1 year (IKDC, Lysholm, Tegner, ICRS, modified Cincinnati) | |
| Open vs. arthroscopic ACI | | | | |
| Harris 2010 ⁴⁴ | Clinical outcome | 1 comparative cohort | IKDC (objective) results significantly better for arthroscopic group at 1 year [effect size 0.58 (SE 0.21)] but no significant difference at 5 years | |

TABLE 54 Results and conclusions of systematic reviews (continued)

| Review | Outcome | No. of studies | Result of meta-analysis/review | Comments |
|---------------------------------|--------------------------------------|----------------|--|----------|
| ACI vs. mosaicplasty/OAT | | | | |
| Bekkers 2009 ⁵⁰ | Clinical outcome | 1 RCT | No significant difference in modified Cincinnati good–excellent score (> 55) at 19 months (ACI 88%, mosaicplasty 69%) | |
| | Macroscopic/histological outcome | 1 RCT | Significant difference in modified Cincinnati good–excellent score (> 55) at 12 months for medial femur (ACI 88%, mosaicplasty 73%, $p = 0.032$) but not lateral femur or patella ICRS macroscopic grading significantly better with ACI at 12 months (excellent–good ACI 82%, mosaicplasty 34%; $p < 0.01$) Only biopsies from ACI group ($n = 7$ predominantly hyaline, $n = 7$ mixed hyaline and fibrocartilage, $n = 5$ predominantly fibrocartilage) | |
| Harris 2010 ⁴⁴ | Clinical outcome | 1 RCT, 1 CCT | 1 RCT: no significant difference in Lysholm score after 1 year 1 CCT: Lysholm score significantly better after 1 year for mosaicplasty but no significant difference at 2 years | |
| Magnussen 2008 ⁵² | Clinical outcome | 1 RCT, 1 CCT | 1 RCT: no significant difference in modified Cincinnati > 55 (ACI 88%, OAT 69%; $p = 0.27$) (1 year) 1 CCT: significantly better Lysholm scores with OAT [P-ACI 67 (SD 8), OAT 74 (SD 6); $p < 0.05$]; no significant difference in Tegner or Meyers scores (2 years) | |
| | Arthroscopic/histological assessment | 1 RCT, 1 CCT | 1 RCT: per cent with CRA 8–12 significantly better with ACI (ACI 82%, OAT 34%; $p < 0.01$); 74% of ACI patients with hyaline- or mixed hyaline/fibrocartilage-like (not reported for OAT) (1 year) 1 CCT: OAT patients with hyaline cartilage not integrated into surrounding cartilage; P-ACI specimens with mainly fibrocartilage, focalised areas or hyaline-like cartilage (2 years) | |
| | Subgroups | 1 RCT, 1 CCT | 1 RCT: significantly more with modified Cincinnati score > 55 with ACI of patients with femoral condyle lesions only (ACI 88%, OAT 74%; $p = 0.03$) (1 year) | |
| | Complications | 1 RCT, 1 CCT | 1 RCT: 7 poor results, all in OAT group (1 year) Arthofibrosis ACI 0–15%, OAT 7.1–15% (up to 2 years) OAT group only: Superficial wound infection 2.4–5%, DVT 2.4%, postoperative haemarthrosis (10%) | |

continued

TABLE 54 Results and conclusions of systematic reviews (continued)

| Review | Outcome | No. of studies | Result of meta-analysis/review | Comments |
|-------------------------------|----------------------|----------------|--|--|
| Nakamura 2009 ⁵⁴ | Clinical outcome | 2 RCTs, 1 CCT | 1 RCT: modified Cincinnati score significantly better for ACI than OAT in the medial femoral condyle (19 months) 1 RCT: no significant difference in Lysholm scores, IKDC (36 months); 1 CCT significantly better Lysholm scores with OAT (2 years) | |
| Naveen 2012 ⁴⁹ | Clinical outcome | 2 RCTs, 1 CCT | 1 CCT: no difference in clinical scores, improvement with ACI lagged behind improvement with mosaicplasty (2 years) 1 RCT: 88% good and excellent after ACT, 69% after mosaicplasty ($p < 0.05$, 19 months) 1 RCT: complete recovery in 68% after ACI, 88% after mosaicplasty (but difference presumably non-significant as treatments are considered equivalent, 36 months) | Actual values not reported for any of the outcomes |
| | Histological outcome | 1 RCT, 1 CCT | 1 CCT: fibrocartilaginous defect filling with ACI, no visible changes in tissue after mosaicplasty (24 months) 1 RCT: 82% good or excellent after ACI, 34% after mosaicplasty (19 months) | |
| Vasiliadis 2010 ⁴⁵ | Clinical outcome | 2 RCTs, 1 CCT | 1 CCT: significantly better recovery (Lysholm) with mosaicplasty than P-ACI (up to 2 years; $p = 0.012$), no significant difference in Tegner or Meyers score 1 RCT: no significant difference in Lysholm score (10 months) 1 RCT: no significant overall difference between P-ACI/C-ACI and mosaicplasty, but ACI significantly better for medial femoral condyle lesions at 12 months (88% good or excellent results vs. 74% for mosaicplasty; $p = 0.032$) | |
| (Cochrane review) | Satisfactory outcome | 2 RCTs, 1 CCT | Meta-analysis showed no significant difference (risk ratio 1.02, 95% CI 0.81 to 1.28; $p = NS$), significant heterogeneity | |
| | Histology | 2 RCTs | 1 RCT: 82% good or excellent after ACI, 34% after mosaicplasty (12 months; $p < 0.01$) – fibrous tissue between grafts in four mosaicplasty patients, plugs disintegrated in three patients; in one ACI patient, with mixed hyaline–fibrohyaline repair tissue, ongoing maturation of repair tissue to hyaline-like tissue was seen 2 years postoperatively 1 RCT: only short-term results – fibrocartilage in central and superficial layers and hyaline cartilage only in deep-layer areas 6 months after ACI' good quality of cartilage of transplanted plugs (but > 50% of biopsies taken at 3 months) | |
| | Complications | 1 CCT | No significant differences in complication rates | |

TABLE 54 Results and conclusions of systematic reviews (continued)

| Review | Outcome | No. of studies | Result of meta-analysis/review | Comments |
|----------------------------|----------------------------------|----------------|---|----------|
| Vavken 2010 ⁵⁶ | Clinical outcome | 2 RCTs, 1 CCT | 1 RCT: no significant difference 1 RCT: complete recovery in 68% after ACI, 88% after mosaicplasty 1 RCT: 88% good and excellent after ACT, 69% after mosaicplasty ($p < 0.05$, 19 months) | |
| | Histology | 2 RCTs | 1 RCT: fibrocartilaginous filling after ACI, no visible changes in tissue after OAT (2 years) 1 RCT: 82% good or excellent after ACI, 34% after mosaicplasty (19 months) | |
| | Complications | 1 RCT, 1 CCT | 1 RCT: four failed treatments with OAT 1 CCT: gaps between plugs and adjacent tissue in all second look arthroscopies | |
| ACI vs. MF | | | | |
| Bekkers 2009 ⁵⁰ | Clinical outcome | 2 RCTs | 1 RCT: CCI vs. MF no significant difference in KOOS at 18 months 1 RCT: ACI vs. MF no significant difference in Lysholm, VAS or Tegner scores at 5 years; SF-36 physical functioning significantly better with MF at 2 years ($p = 0.01$), no significant difference at 5 years | |
| | Macroscopic/histological outcome | 2 RCTs | 1 RCT: significantly higher histomorphometric score with CCI than with MF ($p = 0.003$), as well as significantly higher histology assessment score ($p = 0.012$) 1 RCT: no significant difference in ICRS macroscopic grading between ACI and MF at 2 years; histology ($n = 67$): hyaline ACI 19%, MF 11%; hyaline/fibrocartilage ACI 31%, MF 17%; fibrocartilage ACI 34%, MF 57%; no tissue ACI 16%, MF 15% | |
| | Subgroups | 1 RCT | Better clinical outcomes for both groups for age < 30 years ($p = 0.007$ at 2 years and $p = 0.013$ at 5 years) Lesions < 4 cm ² showed better clinical results in the MF group ($p < 0.003$) | |
| Goyal 2013 ⁴⁸ | Clinical outcome | 7 comparative | Numerical data only reported for MF, no results reported for comparison with ACI | |
| Harris 2010 ⁴⁴ | Clinical outcome | 6 RCTs, 1 CCT | Participants in 3/7 studies had significantly better clinical scores after 1–5 years with ACI than with MF (effect sizes for Lysholm, Tegner, ICRS, KOOS 0.66–1.52); no significant difference for the rest of the studies (KOOS, Lysholm, SF-36 physical component); 1 RCT had significantly better results on the SF-36 physical component at 2 years for MF (effect size -0.65 for ACI) | |

continued

TABLE 54 Results and conclusions of systematic reviews (continued)

| Review | Outcome | No. of studies | Result of meta-analysis/review | Comments |
|------------------------------|--------------------------------------|----------------|--|---|
| Magnussen 2008 ⁵² | Histological outcome | 1 RCT | 1 RCT had a significant difference in histomorphology and histology score in favour of ACI at 1 year | |
| | Durability | 2 RCTs, 1 CCT | Clinical results for MF tended to plateau or deteriorate at longer FUs, whereas results for ACI tended to improve (three studies); at 5 years, sports activity remained stable in the ACI group but declined in the MF group (1 CCT) | |
| | Clinical outcome | 1 RCT | At 2 years, no significant difference in Lysholm or VAS scores (2 years) | SF-36 physical component significantly better with MF (46 (SD 2 vs. P-ACI 42 (SD 2; $p = 0.01$)) |
| | Arthroscopic/histological assessment | 1 RCT | At 2 years, no significant difference in CRA | No significant difference in percentage with hyaline- or mixed hyaline/fibrocartilage-like (MF 29%, P-ACI 50%; $p = 0.08$) |
| | Subgroups | 1 RCT | At 2 years, patients < 30 years ($p = 0.007$) and patients with Tegner scores > 4 ($p = 0.0005$) had better SF-36 scores in both groups; higher SF-36 scores in MF group associated with lesions < 4 cm ² ($p = 0.003$) | |
| Mithöfer 2009 ⁵³ | Complications | 1 RCT | P-ACI: 25% tissue hypertrophy | MF: 7.5% tissue hypertrophy, 2.5% arthrofibrosis |
| | Clinical outcome | 1 RCT | Higher increases in KOOS (sports and recreation) with ACI than MF | |
| | Histology | 1 RCT | Significantly better histological assessment ($p < 0.05$) and histomorphometric scores, including higher proteoglycan content, higher type II collagen content, and more normal chondrocyte morphology ($p < 0.01$) after characterised ACI compared with MF at 12–18 months | |
| Nakamura 2009 ⁵⁴ | Clinical outcome | 2 RCTs, 1 CCT | 1 RCT: no significant difference in Lysholm, Tegner, VAS scores; SF-36 physical component significantly better with MF (2 years); no significant difference in any of the scores at 5 years | |
| | | | 1 RCT: no significant difference in KOOS (18 months) | |
| | | | 1 CCT: significantly better IKDC scores with ACI at 5 years | |
| | Histology | 2 RCTs | No significant difference in 1 RCT (2 years), better result for ACI in 1 RCT (18 months) | |

TABLE 54 Results and conclusions of systematic reviews (continued)

| Review | Outcome | No. of studies | Result of meta-analysis/review | Comments |
|-------------------------------|--|---|--|--|
| Naveen 2012 ⁴⁹ | Clinical outcome | 3 RCTs, 2 comparative cohort | No significant difference in clinical scores in 1 RCT and 1 comparative cohort (2–5 years), ACI better in 1 RCT and 1 comparative cohort (12 months to 5 years), 1 RCT no significant difference at 18 months, but ACI significantly better at 36 months | Actual values not reported for any of the outcomes |
| | Histological outcome | 2 RCTs | No significant difference in 1 RCT (2 years), better result for ACI in 1 RCT (18 months) | |
| | QoL (SF-36) | 2 RCTs | No significant difference in 1 RCT (5 years), better result for MF in 1 RCT (2 years) | |
| Negrin 2013 ⁵⁵ | Clinical outcome | 4 RCTs | At 1 year, SMD 1.05 (95% CI –1.35 to 3.45); $p = \text{NS}$; heterogeneity, $p < 0.0001$ | |
| | | 4 RCTs, 1 comparative cohort | At 2 years, SMD 0.38 (95% CI –0.13 to 0.90); $p = \text{NS}$; heterogeneity, $p = 0.0008$ | |
| | | 2 RCTs, 1 comparative cohort | At 5 years, SMD 0.28 (95% CI –0.23 to 0.79); $p = \text{NS}$; heterogeneity, $p = 0.0143$ | |
| | Subgroups: second and third generation ACI | 3 RCTs | At 1 year, SMD 2.22 (95% CI 1.01 to 3.42); $p < 0.05$; heterogeneity $p = 0.0003$ | |
| | | 3 RCTs, 1 comparative cohort | At 2 years, SMD 0.56 (95% CI 0.30 to 0.82), $p < 0.05$; heterogeneity, $p = \text{NS}$ | |
| | | 1 RCT, 1 comparative cohort | At 5 years, SMD 0.51 (95% CI 0.21 to 0.80), $p < 0.05$; heterogeneity, $p = \text{NS}$ | |
| | Treatment failure | 4 RCTs, 2 comparative cohort | Overall, 21 treatment failures with MF vs. 16 with ACI | |
| | Histology | 2 RCTs | 1 RCT: no significant difference between ACI and MF, but ACI biopsy specimens tended to have a more hyaline-like appearance 2 years postoperatively | |
| | | 1 RCT: clear morphological superiority of cartilaginous tissue after ACI; MF resulted in significantly lower histological scores for type II collagen and matrix proteoglycan content | | |
| Vasiliadis 2010 ⁴⁵ | Clinical outcome | 3 RCTs | 1 RCT: no significant difference between P-ACI and MF (5 years) in Lysholm or Tegner scores or VAS; SF-36 significantly better with MF at 2 years but no significant difference at 5 years 1 RCT: MACI more improvement in Lysholm and Tegner scores than MF but unclear if the difference was significant (12 months) 1 RCT: no significant difference in modified KOOS at 18 months; CCI slightly better at 36 months ($p = 0.05$), slower recovery with CCI, but no significant difference in function at 2 years | |

continued

TABLE 54 Results and conclusions of systematic reviews (continued)

| Review | Outcome | No. of studies | Result of meta-analysis/review | Comments |
|-------------------------------|----------------------|----------------------|--|--|
| Vavken 2010 ⁵⁶ | Histology | 2 RCTs | 1 RCT (12 months): significantly better histomorphogenic score ($p = 0.003$) and better mean histology score ($p = 0.012$) with CCI than MF – obvious cartilaginous restoration after chondrocyte implantation, repair scar tissue after MF | |
| | | | 1 RCT: 71.4% poor-quality repair tissue with MF vs. 50% with P-ACI (2 years), but no statistically significant difference; no association between histological quality and clinical outcomes at 2 and 5 years, but the worse the image at 2 years, the bigger the risk of failure up to 5 years ($p = 0.02$) | |
| | Complications | 2 RCTs | 1 RCT (2 years): 25% debridement due to graft hypertrophy with P-ACI, 10% with MF, 23% in each group had a failure (one in each group a total arthroplasty) | |
| | | | 1 RCT (3 years): similar complication rates with CCI and MF – 2/57 failures with CCI and 7/61 with MF | |
| | Clinical outcome | 3 RCTs | 1 RCT (12 months): significantly better results with ACI than MF; 1 RCT no significant difference in clinical scores (2 and 5 years), 1 RCT no significant difference at 18 months but ACI significantly better at 36 months; 1 RCT SF-36 significantly better with MF than ACI at 2 years but not at 5 years | |
| | Histology | 2 RCTs | 1 RCT: no significant difference (2 years), 1 RCT: better results for ACI at 18 months | |
| | Complications | 2 RCTs | 1 RCT: nine failures in each group; 25% debridement with ACI and 10% with MF (after 5 years) | |
| | | | 1 RCT: 25% cartilage hypertrophy with ACI, 13% with MF, 67% and 59% AEs with ACI and MF (9% and 13% serious) | |
| ACI vs. BMSC | | | | |
| Naveen 2012 ⁴⁹ | Clinical outcome | 1 comparative cohort | Significantly better clinical outcomes for BMSC than ACI (2 years) | Actual values not reported for any of the outcomes |
| | Histological outcome | 1 comparative cohort | Comparison not possible: histological results presented only for BMSC, not ACI | |
| ACI vs. abrasionplasty | | | | |
| Naveen 2012 ⁴⁹ | Clinical outcome | 1 RCT | Significantly better clinical outcomes for ACI (12 months) | Actual values not reported for any of the outcomes |
| Magnussen 2008 ⁵² | Clinical outcome | 1 RCT | At 1 year, significantly better clinical scores in MACI than abrasion group [Lysholm MACI 86 (SD 9), abrasion 74 (SD 11) ($p = 0.001$); IKDC MACI 76 (SD 13), abrasion 68 (SD 10) ($p < 0.05$); Tegner MACI 5.9 (SD 0.8), abrasion 4.2 (SD 11) ($p < 0.01$)] | |

TABLE 54 Results and conclusions of systematic reviews (continued)

| Review | Outcome | No. of studies | Result of meta-analysis/review | Comments |
|---------------------------|-----------------------------------|----------------|---|----------|
| | Histological outcome | 1 RCT | At 1 year, histology on four samples (presumably MACI): evidence of hyaline-like cartilage; fibroblast-like cells in two | |
| | Complications | 1 RCT | 24% reactive synovitis in MACI group | |
| Nakamura 200953 | Clinical outcome | 1 RCT | At 1 year, significantly better Lysholm and IKDC scores with MACI than abrasion | |
| Vasiliadis 201044 | Clinical outcome – Lysholm scores | 1 RCT | ACI significantly better than abrasion at 1 year ($p < 0.001$ for improvement in Lysholm scores, 72% with ACI vs. 40% with abrasion good or excellent results; $p < 0.01$ for difference in Tegner score), IKDC subjective score also significantly better for ACI | |
| Vavken 2010 ⁵⁶ | Clinical outcome | 1 RCT | Significantly better clinical outcomes for ACI (12 months) | |

AKS, arthroscopic knee surgery; BMSC, bone marrow-derived mesenchymal stem cell; C-ACI, collagen-based ACI; CRA, Cartilage Repair Assessment; FU, follow-up; NS, not significant.

TABLE 55 Systematic review conclusions

| Study | Conclusions | Recommendations | Comments |
|----------------------------------|---|---|---|
| Bekkers 2009⁵⁰ | Clinical outcomes: all trials showed an improvement from clinical baseline scores, regardless of treatment; lesion size, activity level and patient age are factors that should be considered in selecting treatment of articular cartilage lesions of the knee | Practice: small chondral and osteochondral lesions (< 1 cm ²) should preferably be treated by MF or single-plug OAT; for larger lesions (> 4 cm ²) MF has been associated with limited effectiveness; for larger lesions, OAT and ACI are both good treatment options Research: patients in trials should be stratified based on BMI, defect location, and post-debridement defect size; outcomes should be reported after at least 2 years of FU using biopsy, MRI and validated clinical outcome tools, including assessment of activity level | |
| Goyal 2013⁴⁷ | General: C-ACI is marginally more effective than P-ACI, with evidence limited to a FU period of 2 years; MACI gives comparable results to P-ACI or C-ACI (evidence from studies with a short duration of FU, with a small sample size and medium-sized defects in a younger age group) | Practice: NR Research: multicentre RCTs with adequate sample size needed of second-and third-generation ACI vs. first-generation ACI; cohort studies of long-term effects (10 years) needed | |
| Goyal 2013⁴⁸ | Only refers to MF | | Publications including the same study populations counted as separate studies |

continued

TABLE 55 Systematic review conclusions (continued)

| Study | Conclusions | Recommendations | Comments |
|------------------------------------|---|---|---|
| Harris 2010⁴⁴ | <p>General: studies were very heterogeneous and had important quality limitations</p> <p>Clinical outcomes: intermediate-term clinical outcomes after ACI tended to be better than after MF; difference compared with osteochondral autograft unclear; no significant differences in clinical outcomes between first- and second-generation ACI</p> <p>Histology: ACI may provide a more durable repair tissue than MF</p> <p>Modifying factors: outcomes tended to be better for younger patients (< 30/35 years), more active patients, patients with shorter symptom duration and patients who had not had a previous failed surgical intervention; possibly better results for smaller lesions and better effects of ACI than other techniques for larger lesions</p> <p>Complications: graft hypertrophy highest with ACI-P (22%), lower with other methods (4–7%); reported 'failure' rates slightly lower with ACI (2.8%) than with MF (3.7%) or mosaicplasty (7.1%)</p> | <p>Practice: ACI may be the best option for large defects in young, active patients with a short duration of symptoms and no previous cartilage surgery; MF is indicated for smaller defects in young, active patients; osteochondral autograft may provide a more rapid improvement in terms of clinical outcome, but is limited by donor site morbidity</p> <p>Research: higher-quality studies needed, with the following characteristics: proper and transparent patient enrolment with clearly stated inclusion and exclusion criteria; proper independently performed randomisation techniques; no concurrent surgical interventions (anterior cruciate ligament reconstruction, realignment osteotomy, meniscal surgery, etc.); consistent surgical technique; longer clinical FU with an independent observer; use of validated, responsive and reliable outcome measures; clear reporting of data with a statement of both clinical relevance and significance</p> | Publications including the same study populations counted as separate studies |
| Kon 2009⁵¹ | <p>Clinical outcomes: matrix-assisted second-generation ACI is a promising technique for the treatment of isolated chondral defects; good clinical results were reported by all products, but FUs were short and quality levels of studies were low</p> | <p>Practice: NR</p> <p>Research: high-quality, long-term RCTs are needed</p> | |
| Magnussen 2008⁵² | <p>General: FU relatively short, heterogeneous outcome measures</p> <p>Clinical outcomes: all trials revealed short-term improvement in all clinical scores with every treatment method evaluated (ACI, MACI, OAT, MF, abrasion)</p> | <p>Practice: MF is ideal first-line treatment for small stage III or IV articular cartilage defects; more complex surgery needed for larger lesions (larger than 2–4 cm²)</p> <p>Research: large multicentre trial needed comparing ACI, MACI, OAT, MF, simple debridement and a non-operative control; trial should use validated patient-orientated clinical outcome measures, e.g. the KOOS, the WOMAC®, SF-36 score, or the IKDC score, with FU at 5 and 10 years</p> | |

TABLE 55 Systematic review conclusions (continued)

| Study | Conclusions | Recommendations | Comments |
|------------------------------------|--|--|---|
| Mithöfer 2009 ⁵³ | Return to sports: return to sports was possible in 73% overall, with highest return rates after osteochondral autograft transplantation; time to return to sports was between 7 and 18 months (longest with ACI); initial return to sports at the pre-injury level was possible in 68% and did not significantly vary between surgical techniques; continued sports participation at the pre-injury level was possible in 65%, with the best durability after ACI; several factors affected the ability to return to sport after ACI: athlete's age (better at younger age), preoperative duration of symptoms (better with shorter duration) | Practice: NR Research: systematic research is needed to explain lack of return to sports and unsustained sports participation in some patients; prospective long-term studies are needed to determine if articular cartilage repair in athletes can influence the high incidence of OA associated with high impact sports | |
| Nakamura 2009 ⁵⁴ | General: studies were of limited quality; there is insufficient evidence from the included studies to say whether or not cell-based therapy is superior to other treatment strategies in articular cartilage lesions of the knee | Practice: NR Research: high-quality RCTs with long-term FU are needed | Publications including the same study populations were counted as separate studies |
| Naveen 2012 ⁴⁹ | Clinical outcomes: there is heterogeneity and inconsistency between studies; it is unclear to what extent any differences between treatments in clinical outcomes are clinically important Histology: ACI is associated with superior structural regeneration of cartilage tissue compared with other methods (but reported by only 6/17 studies) | Practice: NR Research: studies of long-term effects needed | Stated that non-RCTs were excluded, but not all of the included trials were RCTs; publications including the same study populations counted as separate studies |
| Negrin 2013 ⁵⁵ | Clinical outcomes: the meta-analyses (of all forms of ACI vs. MF or only second- and third-generation ACI) did not reveal any clinically relevant superiority of ACI over MF – results converged over time; decision-making must take patient objectives, physical demands and patient- and defect-specific factors into consideration (e.g. MF has worse outcomes with defect sizes > 4 cm ²) | Practice: NR Research: large, well-designed, long term multicentre studies needed | |

continued

TABLE 55 Systematic review conclusions (continued)

| Study | Conclusions | Recommendations | Comments |
|--------------------------------------|--|--|----------|
| Vasiliadis 2010 ⁴⁵ | <p>General: studies are of poor quality, heterogeneity regarding techniques followed and populations studied</p> <p>Clinical outcomes: body of evidence does not suggest superiority of ACI over other techniques; complication rates were comparable between interventions except from an increased rate of graft hypertrophies after P-ACI; ACI is an effective treatment for full-thickness chondral defects of the knee, providing an improvement of clinical outcomes</p> | <p>Practice: there is insufficient evidence to conclude whether or not autologous cartilage implantation is superior to other treatment strategies for treating full-thickness articular cartilage defects in the knee</p> <p>Research: there is a need for more high-quality RCTs and uniformity of their reported outcomes; more studies should be done on maturation process of finally formed repair tissue and appropriate rehabilitation programmes for the different techniques; more information and research is needed to compare chondrocyte techniques with conservative treatment, such as intensive physiotherapy; further information is needed on the relationship between clinical, histological and radiological outcomes, and the most appropriate measure of functional outcomes that relate to a generic measure of health-related QoL</p> | |
| Vavken 2010 ⁵⁶ | <p>General: rather low overall quality of studies, including high attrition rates and small sample sizes</p> <p>Clinical outcomes: some evidence for better clinical outcomes with ACI compared with OAT and equivalent outcomes with MF in studies with higher validity; higher-quality repair tissue with ACI compared with other procedures; unclear if statistical significance corresponds to real clinical significance</p> | <p>Practice: no clear recommendation regarding ACI vs. other treatments possible</p> <p>Research: evolution of techniques needs to be taken into account; further high-quality studies needed</p> | |

FU, follow-up; NR, not reported; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index.

Appendix 7 Excluded studies: survival analysis

| Author ID/year | Reason |
|---------------------------------------|---|
| Adachi 2014 ¹⁹³ | The procedure described seems to be about implantation of cartilage-like tissue rather than chondrocytes |
| Bert 2015 ¹²⁹ | Editorial and opinion piece with no primary data |
| Bae 2013 ¹³⁴ | All had Kellgren–Lawrence score of 3. EMA MACI SPC excluded such patients Patients had OA and mean age 62.1 years so would not be considered for ACI |
| Behery 2014 ⁸⁴ | Systematic review. Used only for checking completeness of our search retrieval. Six studies with 50 patients in case series |
| Brix 2012 ²⁰⁸ | The 8 years' details are too sparse to be of much use. It is only an abstract and we have other much better ACI data |
| Briggs 2013 (abstract) ²⁰⁹ | Mean FU only 4 years. No data on subgroup with longer FU |
| Ebert 2013 ²¹⁰ | Has patients from Ebert 2011 ²¹¹ |
| Ebert 2011 ²¹¹ | Case series Matrix-applied autologous chondrocyte implantation (MACI) Excluded because almost half had concomitant procedures. No FU beyond 5 years |
| Ebert 2013 ²¹² | Includes too many patients having concomitant procedures |
| Filardo 2012 ²¹³ | Second-generation ACI Hyalograft C |
| Filardo 2013 ¹⁸⁹ | Hyaluronan-based scaffold Hyaff 11 (Fidia Advanced Biopolymers Laboratories, Padua, Italy) |
| Filardo 2014 ²¹⁴ | Hyalograft: HYAFF 11 (Fidia Advanced Biopolymers Laboratories) |
| Filardo 2014 ²¹⁵ | Hyalograft C (Fidia Advanced Biopolymers Laboratories) |
| Gobbi 2014 ¹²⁷ | Excluded because large proportion had concomitant surgery such as meniscectomy, cruciate ligament repairs |
| Gooding 2006 ²³ | Only 2 years' FU |
| Gudas 2012 ¹²⁸ | Mosaic-type osteochondral autologous transplantation (OAT) and MF, but only 30 patients in each arm |
| Health Quality Ontario ²¹⁶ | Not about MF or ACI |
| Kon 2009 ⁷³ | Second-generation ACI Hyalograft C |
| Kon 2011 ²¹⁷ | Biocompatible and biodegradable hyaluronan-based scaffold (Hyalograft C) |
| Kon 2011 ⁷⁴ | Arthroscopic Hyalograft C technique |
| Kon 2011 ²¹⁸ | Second-generation ACI (Hyalograft C) |
| Kon 2009 ²¹⁹ | Systematic review |
| Kreuz 2006 ²²⁰ | FU too short |
| McNickle 2009 ²²¹ | FU too short |
| Minas 2012 ²²² | FU only 12 months |
| Mithöfer 2009 ⁸² | Systematic review: mentions only five studies with FU > 5 years |
| Mithöfer 2012 ²⁸ | Review: mentions only five studies with FU > 5 years |
| Nawaz 2011 ²²³ | Only an abstract |

| Author ID/year | Reason |
|--------------------------------------|--|
| Ebert 2011 (abstract) ²¹¹ | Case series abstract only; $n = 41$ patients (44 knees; 53 grafts) |
| Negrin 2013 ⁵⁵ | FU 2–5 years |
| Negrin 2012 ¹⁰⁴ | Most studies in meta-analysis had FU only 2 years. Some had 5 years but we have the individual trials |
| Niemeyer 2010 ²²⁴ | FU too short |
| Noyes 2013 ²²⁵ | Review. Checked for studies |
| Oussedik 2015 ²²⁶ | Systematic review. We have all the individual trials that are eligible |
| Rosenberger 2008 ²²⁷ | Mean FU < 5 years and quite a lot had other procedures, such as osteotomy, so pure ACI < 40 patients |
| Salzmann 2013 ¹²¹ | Minimum postoperative FU of 2 years FU time: 4.2 ± 1.8 years |
| Sciarretta 2013 ²²⁸ | 19 patients Polyvinyl alcohol hydrogel implant (Cartiva Synthetic Cartilage Implant, Cartiva, Alpharetta, GA, USA) |
| Scillia 2015 ²²⁹ | Not ACI or MF. Debridement |
| Ulstein 2014 ²³⁰ | MF technique (MF) vs. osteochondral autologous transplantation (OAT mosaicplasty) MF $n = 11$ OAT mosaicplasty $n = 14$ |
| Upmeier 2007 ²³¹ | FU costs Patients had to have been diagnosed with knee cartilage defects and, according to their operation record, treated between 1997 and 2001 with any of the following techniques: ACI, osteochondral allografts or autografts, MF or subchondral drilling, chondroplasty/laser chondroplasty, abrasion arthroplasty, debridement/cartilage shaving (without further information) |
| Wylie 2015 ²³² | Systematic review |
| Zak 2012 ²³³ | Two-step procedure, a biopsy sample was arthroscopically harvested to culture the cells and to seed them on a matrix [MACI (Genzyme, Cambridge, MA, USA), 15 patients; Hyalograft C (Fidia Advanced Biomaterials, Abano Terme, Italy), 44 patients; CaReS (Arthro Kinetics Biotechnology GmbH, Krems, Austria), 11 patients] |

FU, follow-up; SPC, Summary of Product Characteristics.

Appendix 8 Included long-term studies: data extraction and quality assessment

Asik 2008

| Asik 2008 ¹⁰⁶ | Data |
|--|--|
| Title | The Microfracture Technique for the Treatment of Full-Thickness Articular Cartilage Lesions of the Knee: Midterm Results |
| Type of study | Cohort study (pre–post) |
| | Eligibility criteria reported |
| Quality of study NIH | Good |
| No. of patients | 90 |
| Population | 34.5 years (range 20–58 years) |
| | 47.8% male |
| | Reason for injury not reported |
| Intervention | MF |
| Duration of injury? | NR |
| Previous attempts at repair? | NR |
| Size of defect in cm ² ; depth or severity if given | Mean not reported |
| | Reports <i>n</i> with < 2 cm and ≥ 2 cm (see subgroup results) |
| Duration of FU | 1.5, 3, 6, 12 months and last visit |
| | Mean 68 months (range 24–108 months) |
| Survival curve provided? | No |
| Results | |
| Lysholm score, mean (SD) [range] | Preop.: 52.4 (6.2) [38–70] |
| | Last FU: 84.6 (7.8) [68–100] |
| | Change: 30.4 (4.2) |
| | $p < 0.0001$ |
| Tegner activity scale scores, mean (SD) [range] | Preop.: 2.6 (1.5) [2–5] |
| | Last FU: 5.2 (1.3) [4–9] |
| | Change 2.6 (0.8) |
| | $p < 0.0001$ |
| Oxford knee questionnaire, mean (SD) [range] | Preop.: 23.1 (4.8) [12–30] |
| | Last FU: 44.8 (5.7) [24–48] |
| | Change: 21.7 (3.8) |
| | $p < 0.0001$ |
| Subgroup data given? | |

| Asik 2008 ¹⁰⁶ | Data |
|---|---|
| Lysholm score, mean (SD) | <p>Age:</p> <ul style="list-style-type: none"> • < 35 years, $n = 42$, 36.2 (5.8) • ≥ 35 years, $n = 48$, 24.3 (6.1) • $p < 0.001$ <p>Size of defect:</p> <ul style="list-style-type: none"> • < 2 cm², $n = 68$, 37.4 (5.9) • ≥ 2 cm², $n = 22$, 26.9 (4.7) • $p < 0.001$ <p>Location of defect:</p> <ul style="list-style-type: none"> • weight-bearing surface $n = 42$, 26.8 (5.3) • non-weight-bearing surface, $n = 48$, 37.3 (6.4) • $p < 0.001$ <p>BMI:</p> <ul style="list-style-type: none"> • < 25 kg/m², $n = 52$, 38.2 (5.4) • ≥ 25 kg/m², $n = 38$, 26.2 (4.8) • $p < 0.001$ |
| Tegner activity scale scores, mean (SD) | <p>Age:</p> <ul style="list-style-type: none"> • < 35 years, $n = 42$, 2.6 (0.8) • ≥ 35 years, $n = 48$, 2.1 (0.4) • $p < 0.001$ <p>Size of defect:</p> <ul style="list-style-type: none"> • < 2 cm², $n = 68$, 2.8 (0.6) • ≥ 2 cm², $n = 22$, 2.0 (0.4) • $p < 0.001$ <p>Location of defect:</p> <ul style="list-style-type: none"> • weight-bearing surface $n = 42$, 2.2 (0.5) • non-weight-bearing surface, $n = 48$, 2.6 (0.6) • $p < 0.001$ <p>BMI:</p> <ul style="list-style-type: none"> • < 25 kg/m², $n = 52$, 2.8 (0.4) • ≥ 25 kg/m², $n = 38$, 2.0 (0.3) • $p < 0.001$ |
| Oxford Knee Questionnaire, mean (SD) | <p>Age:</p> <ul style="list-style-type: none"> • < 35 years, $n = 42$, 21.7 (3.4) • ≥ 35 years, $n = 48$, 16.5 (2.8) • $p < 0.001$ <p>Size of defect:</p> <ul style="list-style-type: none"> • < 2 cm², $n = 68$, 22.2 (3.6) • ≥ 2 cm², $n = 22$, 15.8 (2.8) • $p < 0.001$ |

| Asik 2008 ¹⁰⁶ | Data |
|---|---|
| | <p>Location of defect:</p> <ul style="list-style-type: none"> weight-bearing surface $n = 42$, 16.2 (2.7) non-weight-bearing surface, $n = 48$, 23.2 (2.4) $p < 0.001$ <p>BMI:</p> <ul style="list-style-type: none"> $< 25 \text{ kg/m}^2$, $n = 52$, 22.8 (2.1) $\geq 25 \text{ kg/m}^2$, $n = 38$, 16.3 (2.4) $p < 0.001$ |
| Losses to FU: percentage and reasons if given | <p>Excluded:</p> <p>28 lost to regular FU</p> <p>30 who had undergone a secondary surgical intervention after the index operation (16 ACL ruptures, 13 meniscus ruptures and 1 posterior cruciate ligament rupture)</p> <p>98 because an ACL rupture, meniscal lesion, patellofemoral problems, plica lesion, other location of defect, or more than one location of defect was observed at index operation</p> |
| Any costs given? | No |
| Survival curve | No |
| ACL, anterior cruciate ligament; FU, follow-up; NR, not reported; preop., preoperative. | |

Quality Assessment Tool for Case Series Studies: NIH

| Criteria | Yes | No | Other (CD, NR, NA) |
|---|-----|----|--------------------|
| 1. Was the study question or objective clearly stated? | ✓ | | |
| 2. Was the study population clearly and fully described, including a case definition? | ✓ | | |
| 3. Were the cases consecutive? | | | CD |
| 4. Were the subjects comparable? | ✓ | | |
| 5. Was the intervention clearly described? | ✓ | | |
| 6. Were the outcome measures clearly defined, valid, reliable and implemented consistently across all study participants? | ✓ | | |
| 7. Was the length of FU adequate? | ✓ | | |
| 8. Were the statistical methods well described? | ✓ | | |
| 9. Were the results well described? | ✓ | | |

QUALITY RATING: GOOD

CD, cannot determine; FU, follow-up; NA, not applicable; NR, not reported.

Bentley 2012

| Bentley 2012 ⁷⁸ | Data |
|--|--|
| Title | Minimum 10-year Results of a Prospective Randomised Study of Autologous Chondrocyte Implantation vs. Mosaicplasty for Symptomatic Articular Cartilage Lesions of the Knee |
| Type of study | Long-term results of Bentley 2003 ⁵ RCT of MF vs. mosaicplasty, so only MF arm used here |
| Quality of study | Uncertain risk of bias (Cochrane risk of bias tool) |
| No. of patients | ACI: 58 Mosaicplasty: 42 (data not extracted) |
| Population | Total group mean 31.3 years (range 16–49 years) ACI: 30.9 years (16–49 years) 58% male Reason for injury? ACI: trauma 24 (41%); OCD 14 (24%); chondromalacia patellae 12 (21%); other/unknown: 8 (14%) |
| Intervention | ACI-P or ACI-C |
| Duration of injury? | Mean 7.2 years (range 9 months to 20 years) |
| Previous attempts at repair? | 94 (94%) had previous surgery (no details by study arm) |
| (Do not count debridement and lavage – only previous MF, abrasion, drilling, ACI) | No. of previous repairs: mean 1.5 (range 0–4) Included MF, abrasion, debridement, drilling, and carbon-fibre matrix support prostheses |
| Size of defect in cm ² | ACI 44.1 cm ² (range 10–105 cm ²) |
| Depth or severity if given | |
| Duration of FU | Minimum 10 years (range 10–12 years) |
| Survival curve provided? | Yes |
| Results | |
| Failure | ACI 10/58 (17%) Defined as a clinically poor result, with arthroscopic evidence of failure of the graft, or revision surgery to the defect of any kind |
| Modified Cincinnati rating system | ACI: <i>n</i> = 48 (10 failures excluded) |
| Graded as: | <ul style="list-style-type: none"> ● excellent 28 ● good 7 ● fair 6 ● poor 2 |
| <ul style="list-style-type: none"> ● excellent (> 80 points), good (55–79) ● fair (30–54) ● poor (< 30) | 'Excellent' or 'good' seen as significant improvement, 'fair' as marginally better or unchanged, 'poor' as worse |
| Stanmore/Bentley functional rating system | ACI: |
| Function and pain measure, 5-point scale of pain related to function (0 = no pain with any activity, 4 = pain at rest and severe pain with activity) | <ul style="list-style-type: none"> ● 0: 7 ● 1: 23 ● 2: 3 ● 3: 6 ● 4: 4 |

| Bentley 2012 ⁷⁸ | Data |
|--|--|
| Subgroup data | |
| KM estimates (SE) of per cent failure rates at 5 years according to preoperative factors | Age (years), $p = 0.028$: <ul style="list-style-type: none"> • < 26 ($n = 16$) 0 (-) • 26–35 ($n = 25$) 8 (5) • > 35 ($n = 17$) 12 (8) Gender, $p = 0.87$: <ul style="list-style-type: none"> • male ($n = 33$) 6 (4) • female ($n = 25$) 8 (5) Cause, $p = 0.31$: <ul style="list-style-type: none"> • chondromalacia patellae ($n = 11$) 19 (12) • OCD ($n = 11$) 0 (0) • trauma ($n = 29$) 7 (5) • other/unknown ($n = 7$) 0 (-) Site, $p = 0.81$: <ul style="list-style-type: none"> • lateral femoral condyle ($n = 11$) 9 (9) • medial femoral condyle ($n = 24$) 8 (6) • patella ($n = 20$) 5 (5) • other/unknown ($n = 3$) 0 (-) |
| Losses to FU: percentage and reasons if given | ACI 5 (8.6%) Patients who were lost to FU were included until last review and then withdrawn from the study |
| Any costs given? | No |
| Only for papers with survival curves | |
| Is curve KM? | Yes |
| If not, what is it? | |
| Risk table attached? | No |
| Total events reported? | No |
| HRs, p -value and/or 95% CI, and whether adjusted or not | No |
| FU, follow-up. | |

Cochrane risk of bias

| Bias | Author judgement | Support for judgement |
|---|------------------|---|
| Random sequence generation (selection bias) | Low risk | Computer generated |
| Allocation concealment (selection bias) | Unclear risk | Sequential envelopes, unclear if opaque |
| Blinding of participants and personnel (performance bias) | Unclear risk | No details |
| Blinding of outcome assessment (detection bias) | Unclear risk | No details |
| Incomplete outcome data (attrition bias) | Unclear risk | Some missing data for subjective outcomes for one study arm (not relevant to the review though) |
| Selective reporting (reporting bias) | Unclear risk | No information to judge |
| Other bias | Low risk | |

Beris 2012

| Beris 2012 ¹⁰⁷ | Data |
|----------------------------------|---|
| Title | Treatment of Full-thickness Chondral Defects of the Knee with Autologous Chondrocyte Implantation: A Functional Evaluation with Long-term Follow-up |
| Type of study | Case series |
| Quality of study NIH | Fair |
| No. of patients | 42 (45 knees) |
| Population | Mean age 28.9 years (range 12–47 years) 69% male Reason for injury? <ul style="list-style-type: none"> ● Trauma (38/45 knees) ● OCD (7/45 knees) |
| Intervention | ACI-P |
| Duration of injury? | 28 months |
| Previous attempts at repair? | NR |
| Size of defect | Mean 5.33 cm ² (range 1.8–12 cm ²) |
| Depth or severity if given | All had isolated moderate to large full-thickness (Outerbridge grade III or IV) chondral defects |
| Duration of FU | Mean 96 months (range 62–144 months) Evaluation at 6, 12, 24, 48 months and annually thereafter |
| Survival curve provided? | No |
| Results | |
| Lysholm score, median | Preop.: 56.0 Last FU: 89.0 $p < 0.05$ |
| IKDC | Preop.: 45 Last FU: 69 $p < 0.05$ |
| Tegner activity score | Preop.: 5.5 Last FU: 6.5 $p < 0.05$ |
| ICRS | Preop.: 3.8 Last FU: 2.8 $p < 0.05$ |
| Stanmore functional rating score | Preop.: 3.06 Last FU: 0.94 |
| Pain VAS | Preop.: 7.33 Last FU: 2 $p < 0.05$ Does not appear to be a validated scale |

| Beris 2012 ¹⁰⁷ | Data |
|---|------|
| Subgroup data given? | None |
| Losses to FU: percentage and reasons if given | NA |
| Any costs given? | None |
| Survival curve? | No |

FU, follow-up; NA, not applicable; NR, not reported; preop., preoperative.

Quality Assessment Tool for Case Series Studies: NIH

| Criteria | Yes | No | Other (CD, NR, NA) |
|---|-----|----|--------------------|
| 1. Was the study question or objective clearly stated? | ✓ | | |
| 2. Was the study population clearly and fully described, including a case definition? | ✓ | | |
| 3. Were the cases consecutive? | | ✓ | |
| 4. Were the subjects comparable? | ✓ | | |
| 5. Was the intervention clearly described? | ✓ | | |
| 6. Were the outcome measures clearly defined, valid, reliable and implemented consistently across all study participants? | | | CD |
| 7. Was the length of FU adequate? | ✓ | | |
| 8. Were the statistical methods well described? | ✓ | | |
| 9. Were the results well described? | | ✓ | |

QUALITY RATING: FAIR

CD, cannot determine; FU, follow-up; NA, not applicable; NR, not reported.

Additional comments

Selective reporting of study results

Reports median for Lysholm score, mean or median not stated for other outcomes. No measure of variance.

Bhosale 2009

| Bhosale 2009 ¹⁰⁸ | Data |
|------------------------------|--|
| Title | Midterm to Long-Term Longitudinal Outcome of Autologous Chondrocyte Implantation in the Knee Joint |
| Type of study | Cohort study |
| Quality of study NIH | Good |
| No. of patients | 80 |
| Population | Mean 34.6 (SD 9.1) years 78.8% male Reason for injury not reported |
| Intervention | ACI-P |
| Duration of injury? | NR |
| Previous attempts at repair? | Previous repair (not defined) 70/80 (87.5%) had median of 1 (IQR 1–2) repairs |

| Bhosale 2009 ¹⁰⁸ | Data |
|--|---|
| (Do not count debridement and lavage – only previous MF, abrasion, drilling, ACI) | |
| Size of defect in cm ² | Median defect area 4.1 cm ² (IQR 3.0–6.0 cm ²) |
| Depth or severity if given | Maximum size 20 cm ² |
| Duration of FU | Mean 5 years (range 2.7–9.3 years) |
| Survival curve provided? | No |
| Results | |
| Modified Lysholm score, median IQR | Preop.: 54 (IQR 35.5–68.5) 1 year: 78 (IQR 52–87) Median increase of 24 points |
| Subgroup data given? | <ul style="list-style-type: none"> ● Age ● Gender ● Defect size ● Defect location (lateral femoral condyle; medial femoral condyle; multiple defects; trochlea; other) ● Previous procedures ● Baseline Lysholm score |
| | Regression analysis as potential predictors for change in Lysholm score. Results not extracted |
| Losses to FU: percentage and reasons if given | NA |
| Any costs given? | None |
| Survival curve? | No |
| FU, follow-up; IQR, interquartile range; NA, not applicable; NR, not reported; preop., preoperative. | |

Quality Assessment Tool for Case Series Studies: NIH

| Criteria | Yes | No | Other (CD, NR, NA) |
|---|-----|----|--------------------|
| 1. Was the study question or objective clearly stated? | ✓ | | |
| 2. Was the study population clearly and fully described, including a case definition? | ✓ | | |
| 3. Were the cases consecutive? | ✓ | | |
| 4. Were the subjects comparable? | ✓ | | |
| 5. Was the intervention clearly described? | ✓ | | |
| 6. Were the outcome measures clearly defined, valid, reliable and implemented consistently across all study participants? | ✓ | | |
| 7. Was the length of FU adequate? | ✓ | | |
| 8. Were the statistical methods well described? | ✓ | | |
| 9. Were the results well described? | | ✓ | |
| QUALITY RATING: GOOD | | | |
| CD, cannot determine; FU, follow-up; NA, not applicable; NR, not reported. | | | |
| Additional comments | | | |
| Long-term data measured but not reported. | | | |

Biant 2014

| Biant 2014 ⁷⁹ | Data |
|-----------------------------------|--|
| Title | Long-term Results of Autologous Chondrocyte Implantation in the Knee for Chronic Chondral and Osteochondral Defects |
| Type of study | Case series |
| Quality of study | Good |
| No. of patients | 104 |
| Population | Mean age: 30.2 years (range 15–49 years) 52.9% male Reason for injury? <ul style="list-style-type: none"> ● Trauma: 55 (53%) ● OCD: 17 (16%) ● Chondromalacia patellae: 23 (22%) ● Childhood osteomyelitis: 2 (2%) ● Other/unknown: 7 (7%) |
| Intervention | ACI-P |
| Duration of injury? | Mean 7.8 years |
| Previous attempts at repair? | Previous repair (MF, drilling, mosaicplasty, carbon fibre matrix support prosthesis): 73 (70%) had ≥ 1 previous operation 31 (29.8%) had previous arthroscopic surgery and arthroscopic debridement No. of previous repairs: mean 1.3 (range 0–5) |
| Size of defect in cm ² | 4.78 cm ² (range 1.2–25 cm ²) |
| Depth or severity if given | |
| Duration of FU | Minimum of 10 years (range 10–12 years) Mean 5.7 years' graft failure |
| Survival curve provided? | Yes |
| Results | |
| Graft failure | 27 (26%): all occurred within 8 years Definition: patients who underwent revision surgery of any kind (thereby altering or removing the original graft) or arthroplasty |
| Pain, VAS, 10-point scale | Preop.: 6 Change to last FU: –8.3 (95% CI –10.8 to –5.8) |
| Modified Cincinnati knee score: | Preop.: NR Last FU (intact graft, <i>n</i> = 73): 78 (range 10–100) Change: 53 (95% CI 34 to 71) Excellent: 46 (63%) Good: 18 (24.7%) Fair: 6 (8.2%) Poor: 3 (4.1%) |

| Biant 2014 ⁷⁹ | Data |
|--|--|
| Stanmore/Bentley functional rating system | <p>Preop.: not reported. Assume change score is for patients with an intact graft ($n = 73$)</p> <p>Change to last FU: -2.6 (95% CI -3.7 to -1.5)</p> <p>Score ($n = 73?$), n (%)</p> <ul style="list-style-type: none"> ● 0: 14 (19.2%) ● 1: 38 (52.1%) ● 2: 8 (11%) ● 3: 8 (11%) ● 4: 5 (6.8%) |
| Satisfaction | 98/100 (98%) |
| Patients asked by an independent interviewer if they were satisfied with their ACI surgery and whether or not they would consider undergoing it again if the same symptoms arose in the other knee | |
| Complications | <p>3 (2.9%)</p> <p>2 (1.9%) manipulation under anaesthesia within 8 weeks of surgery because of early postoperative stiffness</p> <p>1 (0.96%) DVT</p> |
| Subgroup data given? | <ul style="list-style-type: none"> ● No prior cartilage surgery, $n = 32$ ● Preop. modified Cincinnati knee score: mean 49 (range 18–94); last FU mean 71 (range 10–100) ● Preop. Stanmore/Bentley score mean 3 (range 1–4); last FU mean 1.5 (range 0–4) ● Preop. mean VAS score 7 (range 1–10); last FU: mean 3.5 (range 0–10) ● Graft failures 6/32 (18.7%) |
| Four were lost to FU, but ITT n used here for proportion | <ul style="list-style-type: none"> ● Prior cartilage repair surgery, $n = 72$; states 73 earlier in report ● Preop. modified Cincinnati knee score mean 42 (range 12–82); last FU: mean 65 (range 10–100) ● Preop. Stanmore/Bentley score mean 3 (range 0–4); last FU: 2 (range 0–4) ● Preop. mean VAS score was 5.5 (range 0.5–10); last FU: mean 3.5 (range 0–10) ● Graft failures: 21/72 (29.2%) |
| Patellar lesions, $n = 36$ | <ul style="list-style-type: none"> ● $n = 36$ ● Preop. mean modified Cincinnati knee score: 40 (range 14–73); last FU ($n = 27$): 79 (range 48–100) ● Excellent: 17 (63%) ● Good: 8 (30%) ● Fair: 2 (7%) ● Preop. mean Stanmore/Bentley score 3 (range 2–4); last FU ($n = 27$): 1.3 (range 0–4) ● Preop. mean VAS score: 6.4 (range 2.5–10); last FU ($n = 27$): 2 (range 0–8) ● Graft failure: 9 (25%) ● The mean time to failure 5.8 years (range 1–8 years) |
| Losses to FU: percentage and reasons if given | 4 (3.8%) |
| Any costs given? | No |
| Survival curve? | No |

FU, follow-up; NR, not reported; postop., postoperative; preop., preoperative.

Quality Assessment Tool for Case Series Studies: NIH

| Criteria | Yes | No | Other (CD, NR, NA) |
|---|-----|----|--------------------|
| 1. Was the study question or objective clearly stated? | ✓ | | |
| 2. Was the study population clearly and fully described, including a case definition? | ✓ | | |
| 3. Were the cases consecutive? | ✓ | | |
| 4. Were the subjects comparable? | ✓ | | |
| 5. Was the intervention clearly described? | ✓ | | |
| 6. Were the outcome measures clearly defined, valid, reliable and implemented consistently across all study participants? | ✓ | | |
| 7. Was the length of FU adequate? | ✓ | | |
| 8. Were the statistical methods well described? | ✓ | | |
| 9. Were the results well described? | ✓ | | |

QUALITY RATING: GOOD

CD, cannot determine; FU, follow-up; NA, not applicable; NR, not reported.

Browne 2005

| Browne 2005 ¹⁰⁹ | Data |
|-----------------------------------|---|
| Title | Clinical Outcome of Autologous Chondrocyte Implantation at 5 Years in US Subjects |
| Type of study | Case series prospective registry from 40 centres |
| Quality of study NIH | Poor |
| No. of patients | 100 |
| Population | Mean 37.0 (SD 9.1) years, range 14–55 years 65% male Reason for injury? <ul style="list-style-type: none"> Acute injury: 58/100 (58%) |
| Intervention | ACI-P |
| Duration of injury? | NR |
| Previous attempts at repair? | <ul style="list-style-type: none"> At least one surgical procedure: 78/100 (78%) At least one cartilage repair procedure: 70/100 (70%) Abrasion/drilling/MF: 36/100 (36%) Fragment reattachment/removal: 1/100 (1%) Osteochondral allograft/autograft: 1/100 (1%) ACI: 1/100 (1%) Meniscus repair/meniscectomy: 32/100 (32%) Ligament repair/reconstruction: 14/100 (14%) Patella alignment: 1/100 (1%) Other: 4/100 (4%) |
| Size of defect in cm ² | <ul style="list-style-type: none"> Mean 4.9 cm² (SD 3.8, range 0.84–23.54 cm²) < 2.0 cm²: 15/100 (15%) 2.0 to < 4.0 cm²: 38/100 (38%) 4.0 to < 6.0 cm²: 17/100 (17%) ≥ 6.0 cm²: 30/100 (30%) Multiple defects: 15/100 (15%) |
| Depth or severity if given | |

| Browne 2005 ¹⁰⁹ | Data |
|--|---|
| Duration of FU | 5 years |
| Survival curve provided? | No |
| Results | |
| Overall condition score, mean (SD), $n = 87$, modified Cincinnati knee rating system | Preop.: 3.2 (1.5) 5-year FU: 5.8 (2.8) Change: 2.6 (3.2); $p < 0.0001$ (95% CI 1.9 to 3.2) |
| Pain mean (SD), $n = 86$ | Preop.: 3.1 (2.2) |
| Patient rated measure (6 point scale 0–10), unlikely validated | 5-year FU: 5.5 (3.2) Change: 2.3 (3.7); $p < 0.0001$ (95% CI 1.5 to 3.1) Preop.: 4.1 (2.7) 5-year FU: 6.1 (3.1) Change: 2.0 (3.8); $p < 0.0001$ (95% CI 1.2 to 2.8) |
| Proportion in response sets | Swelling mean (SD), $n = 85$ Improved: 62/100 (62%) No change: 6/100 (6%) Worsened: 19/100 (19%) Definitions not provided; states 'additional examination' |
| Failure: Cases in which a patient needed an operation after ACI that necessitated the removal of the graft, confirmed a loss of defect fill or violated the subchondral bone (e.g. abrasion chondroplasty, MF, drilling, unicompartmental knee replacement, TKR) | 13/100 (13%) |
| Complications | Joint infections, $n = 0$ Arterial injuries, $n = 0$ Nerve injuries, $n = 0$ DVT, $n = 1$ Reflex sympathetic dystrophy, $n = 1$ Closed manipulation under anaesthesia, $n = 2$ |
| Subgroup data given? | Men ($n = 65$) vs. women ($n = 35$): 2.4 vs. 2.8 |
| Modified Cincinnati knee rating system | Concurrent procedures ($n = 21$) vs. no concurrent procedures ($n = 79$): 2.5 vs. 2.6 |
| Overall condition, change from baseline | |
| Overall condition | Patients rated as improved, $n = 62$ Preop.: 3.0 (1.4) 5-year FU: 7.1 (2.2) Change: 4.1 (2.2); $p < 0.0001$ (95% CI 3.6 to 4.7) |
| Losses to FU: percentage and reasons if given | Unable to collect 5-year FU data on 13 participants Numbers reporting outcomes varied from 62 to 87 |

| Browne 2005 ¹⁰⁹ | Data |
|----------------------------|------|
| Any costs given? | No |
| Survival curve? | No |

FU, follow-up; NR, not reported; preop., preoperative.

Quality Assessment Tool for Case Series Studies: NIH

| Criteria | Yes | No | Other (CD, NR, NA) |
|---|-----|----|--------------------|
| 1. Was the study question or objective clearly stated? | ✓ | | |
| 2. Was the study population clearly and fully described, including a case definition? | ✓ | | |
| 3. Were the cases consecutive? | ✓ | | |
| 4. Were the subjects comparable? | ✓ | | |
| 5. Was the intervention clearly described? | ✓ | | |
| 6. Were the outcome measures clearly defined, valid, reliable and implemented consistently across all study participants? | | ✓ | |
| 7. Was the length of FU adequate? | ✓ | | |
| 8. Were the statistical methods well described? | ✓ | | |
| 9. Were the results well described? | | ✓ | |

QUALITY RATING: POOR

CD, cannot determine; FU, follow-up; NA, not applicable; NR, not reported.

Additional comments (if POOR, please state why)

Proportion 'improved' and 'worsened' not defined.

Gomoll 2014

| Gomoll 2014 ¹¹⁰ | Data |
|----------------------------|---|
| Title | Autologous Chondrocyte Implantation in the Patella: A Multicentre Experience |
| Type of study | Case series. Retrospective analysis, but based on a prospective patient registry from four centres specialising in cartilage repair |
| Quality of study | Poor |
| No. of patients | 110 |
| Population | Additional 23 were lost to FU (FU rate 83%) Age – mean 33 (SD 10.1) years (range 15–55 years) 41.8% male No bilateral ACI included All patients with ACI for patellar defects (including trochlear graft) with at least 4 years' FU were included. Defects outside the patellofemoral compartment were excluded Text discusses differences in population by centres (not data extracted) |

| Gomoll 2014 ¹¹⁰ | Data |
|--|---|
| Intervention | ACI-P (procedure described by Minas <i>et al.</i> , 1999) |
| Duration of injury? | Reported symptoms for mean of 3 years (SD 35 months), range 2–144 months |
| Previous attempts at repair? | Mean of 1.2 previous surgery (range 0–12, SD 1.7) Most common prior procedures were chondroplasty and lateral release |
| Size of defect in cm ² | Mean 5.4 (SD 2.7) cm ² ; range 1–13.2 cm ² |
| Depth or severity if given | 30 (27%) had bipolar disease with an additional trochlear defect; mean size of 4.5 (SD 2.8) cm ² , range 1–13 cm ² 12 distal (11%; type I), 3 lateral (3%; type II), 16 medial (15%; type III) and 79 central/panpatellar defects (72%; type IV) by Pidorianno/Fulkerson classification 82 (75%) of patellar defects and 26 (87%) of trochlear defects were circumferentially shouldered by healthy cartilage (contained) |
| Duration of FU | Mean of 90 (SD 31.7) months, range 48–192 months States data collected yearly intervals; patient-reported outcomes analysed at latest FU |
| Survival curve provided? | No |
| Results, measured at latest FU | |
| SF-12 (QoL) | Physical subscale, <i>n</i> = 89 (81%); baseline: 38.6; last FU 44.1 (<i>p</i> = 0.001) Mental subscale, <i>n</i> = 89 (81%); baseline: 49.7; last FU: 53.5 (<i>p</i> = 0.1) |
| KSS | Knee, <i>n</i> = 44 (40%); baseline: 61.8; last FU: 85.2 (<i>p</i> < 0.001) Function <i>n</i> = 44 (40%); baseline: 58.5; last FU: 72.7 (<i>p</i> < 0.0001) |
| IKDC | <i>n</i> = 65 (60%) Baseline: 40.2 Last FU: 69.4 (<i>p</i> < 0.0001) 86% and 74% of patients demonstrated more than 10 and 20 points of improvement, respectively (considered to exceed the minimal clinically important difference) |
| Modified Cincinnati rating scale, range 2–10 | <i>n</i> = 85 (78%); baseline: 3.2; last FU: 6.2 (<i>p</i> < 0.0001) |
| WOMAC | <i>n</i> = 44 (40%); baseline: 50.4; last FU: 28.6 (<i>p</i> < 0.0001) • 75% of patients exceeded a commonly accepted threshold for MCIDs, with more than a 26% improvement in WOMAC from baseline |
| Satisfaction with procedure | <i>n</i> = 93 (84.5%) |
| Measure used not reported | 84% felt improvement at the time of final FU 86% rated their knee function as good or excellent 92% would choose to undergo ACI again |
| Treatment failure | 9/110 (8.2%). If diagnosed by MRI and/or arthroscopy, with structural failure of the ACI graft in conjunction with pain requiring revision surgery |

| Gomoll 2014 ¹¹⁰ | Data |
|---|--|
| Subgroup data given? | <i>p</i> -values only given, data not extracted States that none of the differences among subgroups reached statistical significance: <ul style="list-style-type: none"> ● polarity (bipolar vs. unipolar) ● containment (contained vs. uncontained; patellar defects only) ● concomitant tibial tuberosity transfer (yes vs. no) ● patellar defect location (lateral, medial, panpatellar) ● defect size (< 4 cm² vs. > 4 cm²) ● sex (male vs. female) |
| Losses to FU: percentage and reasons if given | NA (only those not lost to FU were included) Note that questionnaires were added as they became available and validated, and the start date varied between institutions. Therefore, not all patients answered the same battery of questionnaires |
| Any costs given? | None |
| Survival curve? | No |
| FU, follow-up; MCID, minimum clinically important difference; NA, not applicable; KSS, Knee Society Score; SF-12, Short Form questionnaire-12 items; WOMAC, Western Ontario and McMaster Universities Osteoarthritis Index. | |

Quality Assessment Tool for Case Series Studies: NIH

| Criteria | Yes | No | Other (CD, NR, NA) |
|---|-----|----|--------------------|
| 1. Was the study question or objective clearly stated? | ✓ | | |
| 2. Was the study population clearly and fully described, including a case definition? | ✓ | | |
| 3. Were the cases consecutive? | | ✓ | |
| 4. Were the subjects comparable? | ✓ | | |
| 5. Was the intervention clearly described? | ✓ | | |
| 6. Were the outcome measures clearly defined, valid, reliable and implemented consistently across all study participants? | ✓ | | |
| 7. Was the length of FU adequate? | ✓ | | |
| 8. Were the statistical methods well described? | ✓ | | |
| 9. Were the results well described? | ✓ | | |

QUALITY RATING: GOOD

CD, cannot determine; FU, follow-up; NA, not applicable; NR, not reported.

Additional Comments (if POOR, please state why)

Measures of variance not reported.

Jungmann 2012

| Jungmann 2012 ¹¹¹ | Data |
|--|--|
| Title | Autologous Chondrocyte Implantation for Treatment of Cartilage Defects of the Knee |
| Type of study | Case series. Retrospective analysis of prospective database. Described in paper as a cohort study, level 3 evidence |
| Quality of study | Good |
| No. of patients | 413 |
| Population | Age 34.9 (SD 9.0) years 57.4% male Origins of the cartilage defect: <ul style="list-style-type: none"> • Traumatic, 7.0% • Degenerative, 52.0% • Protracted traumatic-degenerative, 28.3% • Previous OCD or flake fracture, 12.6% |
| Intervention | <ul style="list-style-type: none"> • ACI-P, <i>n</i> = 109 • ACI-C, <i>n</i> = 235 • MACI, <i>n</i> = 69 <p>CellGenix (Freiburg, Germany) for cell suspensions [periosteum patch-covered ACI and Chondro-Gide™ (Geistlich, Wolhusen, Switzerland)-covered ACI] or BioTissue Technologies (Freiburg, Germany) for BioSeed-C (matrix associated) procedure</p> |
| Duration of injury? | NR |
| Previous attempts at repair? | No previous knee surgery: 29.8% MF: 18.6% Pridie drilling: 7.3% ACI: 4.2% Abrasion arthroplasty/debridement: 3.1% Mosaicplasty (OATS): 1.9% Autologous spongiosa graft: 1.7% Retrograde drilling: 0.72% |
| Size of defect | 5.6 (SD 3.0) cm ² |
| Depth or severity if given | |
| Duration of FU | 2–11.8 years FU cut-off was at 5 years 62.5% had a FU at 5 years |
| Survival curve provided? | Yes |
| Results | |
| Revision surgery (treatment failure), <i>n</i> (%) | Treatment failure, represented by need for revision surgery, indicated by: <ul style="list-style-type: none"> • persistent pain at the operated knee joint; • significant loss of function of the operated knee joint; and • clinical findings and/or MRI revealed compatibly pathological changes to confirm symptoms, such as MRI evidence of graft delamination, hypertrophy, severely abnormal signal, insufficient fusion with adjacent cartilage, or secondary transplant defects |

| Jungmann 2012 ¹¹¹ | Data |
|--|--|
| | 88/413 (21.3%) |
| | ACI-P: 34/109 (31.2) |
| | ACI-C: 43/235 (18.3) |
| | MACI: 11/69 (15.9) |
| | Periosteum patch-covered technique [$p=0.031$: odds ratio, 2.4 (BioSeed-C) vs. 2.0 (Chondro-Gide)] increased the risk for the need of re-intervention |
| Time to revision surgery, mean (SD) years | <ul style="list-style-type: none"> ● ACI-P: 1.7 (1.2) ● ACI-C: 1.7 (1.1) ● MACI: 2.4 (1.2) |
| | $p = NS$ |
| Subgroup data: | Age (years) |
| Treatment failure (revision), prognostic factors $n\%$ | <ul style="list-style-type: none"> ● < 30: 24/123 (19.5) ● 30–39: 39/179 (21.8) ● ≥ 40: 25/111 (22.5) |
| Defects related to a trauma within the past 6 months before surgical treatment were considered 'traumatic,' while those associated with a traumatic incident more than 6 months before surgical treatment were considered 'posttraumatic.' | BMI |
| Degenerative' defects were considered those cases in which no trauma could be evaluated. | <ul style="list-style-type: none"> ● < 25: 55/232 (23.7) ● 25–29: 25/149 (16.8) ● ≥ 30: 8/32 (25.0) |
| | No. of defects |
| | <ul style="list-style-type: none"> ● 1: 74/340 (21.8) ● > 1: 14/73 (19.2) |
| | Defect size (cm ³) |
| | <ul style="list-style-type: none"> ● < 3: 12/44 (27.3) ● ≥ 3: 76/369 (20.6) |
| | Cause |
| | <ul style="list-style-type: none"> ● Degenerative: 43/215 (20.0) ● Protracted traumatic-degenerative: 26/117 (22.2) ● OCD, flake fracture: 11/52 (21.2) ● Trauma: 8/29 (27.6) |
| | Gender |
| | <ul style="list-style-type: none"> ● Male: 41/237 (17.3) ● Female: 47/176 (26.7) |
| | Location |
| | <ul style="list-style-type: none"> ● Multiple: 13/68 (19.1) ● Patella: 26/111 (23.4) ● Medial femoral condyle: 36/168 (21.4) ● Lateral femoral condyle: 3/37 (24.3) ● Trochlea: 4/29 (13.8) |
| | Nicotine |
| | <ul style="list-style-type: none"> ● No: 59/298 (19.8) ● Yes: 29/115 (25.2) |

| Jungmann 2012 ¹¹¹ | Data |
|---|---|
| | Parallel treatment |
| | <ul style="list-style-type: none"> Without, 67/306 (21.9) With, 21/107 (19.6) |
| | Previous surgery |
| | <ul style="list-style-type: none"> No, 17/123 (13.8) 1, 41/223 (20.2) > 1, 26/67 (38.8) |
| | Previous treatment |
| | <ul style="list-style-type: none"> No, 54/289 (18.7) Bone marrow stimulation, 28/94 (29.8) Previous transplantation, 5/23 (21.7) Other, 1/7 (14.3) |
| | Female gender ($p = 0.015$; odds ratio 1.7), more than one previous surgery ($p < 0.001$; odds ratio 4.0), and previous BMS ($p = 0.017$; odds ratio 1.9), increased the risk for the need of re-intervention |
| Losses to FU: percentage and reasons if given | NA |
| Any costs given? | No |
| Survival curve? | No |

BMS, bone marrow stimulation; FU, follow-up; NA, not applicable; NR, not reported; NS, not significant.

Quality Assessment Tool for Case Series Studies: NIH

| Criteria | Yes | No | Other (CD, NR, NA) |
|---|-----|----|--------------------|
| 1. Was the study question or objective clearly stated? | ✓ | | |
| 2. Was the study population clearly and fully described, including a case definition? | ✓ | | |
| 3. Were the cases consecutive? | | ✓ | |
| 4. Were the subjects comparable? | ✓ | | |
| 5. Was the intervention clearly described? | ✓ | | |
| 6. Were the outcome measures clearly defined, valid, reliable and implemented consistently across all study participants? | ✓ | | |
| 7. Was the length of FU adequate? | ✓ | | |
| 8. Were the statistical methods well described? | ✓ | | |
| 9. Were the results well described? | ✓ | | |
| QUALITY RATING: GOOD | | | |
| CD, cannot determine; FU, follow-up; NA, not applicable; NR, not reported. | | | |

Knutsen 2007

| Knutsen 2007 ⁶⁷ | Data |
|---|--|
| Title | A Randomised Trial Comparing Autologous Chondrocyte Implantation with MF |
| Type of study | RCT |
| Quality of study | Uncertain risk of bias |
| No. of patients | Total 80 <ul style="list-style-type: none"> • ACI, 40 • MF, 40 |
| Population | Reason for injury: <ul style="list-style-type: none"> • Trauma, 65% • OCD, 28% • Unknown, 7% <p>Baseline characteristics available in online supplement – unable to access</p> |
| Intervention | ACI-P MF |
| Duration of injury? | 36 months |
| Previous attempts at repair? | 74 (93%) had previous knee surgery, including anterior cruciate ligament reconstruction (15), meniscal surgery (14), arthroscopic lavage and debridement (29), Pridie drilling (3), operations for OCD, such as drilling or fixation of a fragment (13) |
| Size of defect | No included defects were deeper than 10 mm |
| Depth or severity if given | |
| Duration of FU | 5 years |
| Survival curve provided? | Yes |
| Results | |
| Failures: | <ul style="list-style-type: none"> • ACI: 9/40 (23%) • MF: 9/40 (23%) |
| Operation considered to have failed if patient needed reoperation because of symptoms due to a lack of healing of the treated defect. The need for shaving or trimming of a lesion was not defined as a failure | Failures occurred at a mean of 26.2 months after ACI and 37.8 months after MF ($p = 0.101$) |
| Median Lysholm score (assume range) | ACI estimated from figure: <ul style="list-style-type: none"> • Baseline: 62 (25–90) • 5 years: 78 (21–100) <p>MF:</p> <ul style="list-style-type: none"> • Baseline: 58 (12–95) • 5 years: 80 (37–100) <p>Difference between groups $p = 0.227$ after adjustment for pretreatment values</p> |

| Knutsen 2007 ⁶⁷ | Data |
|--|---|
| VAS pain scale, median (assume range) | Estimated from figure: ACI <ul style="list-style-type: none"> ● Baseline: 52 (2–100) ● 5 years: 26 (0–100) MF <ul style="list-style-type: none"> ● Baseline: 52 (18–83) ● 5 years: 26 (0–86) Difference between groups $p = 0.278$ after adjustment for pretreatment values |
| SF-36 PCS, median (assume range) | Estimated from figure: ACI <ul style="list-style-type: none"> ● Baseline: 42 (26–58) ● 5 years: 48 (20–65) MF <ul style="list-style-type: none"> ● Baseline: 38 (20–56) ● 5 years: 48 (12–68) Difference between groups $p = 0.054$ after adjustment for pretreatment values |
| Proportion compared with baseline | Less pain: 72% Improvement in Lysholm score: 80% Improvement in SF-36 PCS: 72% |
| Mean Tegner score | ACI <ul style="list-style-type: none"> ● Baseline: 3.28 ● 5 years: 4.05; $p = 0.007$ MF <ul style="list-style-type: none"> ● Baseline: 3.16 ● 5 years: 4.36; $p = 0.002$ Difference between groups $p = 0.323$ after adjustment for pretreatment values |
| Subgroup data | |
| No. of failures by 5 years | Histological grade (no. of knees): no. of failures <ul style="list-style-type: none"> ● 1: ($n = 10$) 0 ● 2: ($n = 16$) 3 ● 3: ($n = 29$) 6 ● 4: ($n = 12$) 3 Younger patients (less than 30 years old) had a better clinical outcome than did older patients ($p = 0.013$) regardless of their treatment group |
| None of the patients with a failure had the best-quality cartilage ($p = 0.001$) | Data not presented, unclear if subgroup defined a priori |
| Losses to FU: percentage and reasons if given | No losses to FU The patients with a failure remained in the study, with their last recorded clinical FU scores before the failure considered to be their final clinical score |
| Any costs given? | No |
| Survival curve? | No |
| FU, follow-up; PCS, physical component score. | |

Cochrane risk of bias score

| Bias | Author judgement | Support for judgement |
|---|------------------|-----------------------|
| Random sequence generation (selection bias) | Unclear risk | NR |
| Allocation concealment (selection bias) | Unclear risk | NR |
| Blinding of participants and personnel (performance bias) | Unclear risk | NR |
| Blinding of outcome assessment (detection bias) | Unclear risk | NR |
| Incomplete outcome data (attrition bias) | Low risk | No losses to FU |
| Selective reporting (reporting bias) | Unclear risk | |
| Other bias | Low risk | |

FU, follow-up; NR, not reported.

Krych 2012

| Krych 2012 ¹¹² | Data |
|-----------------------------------|--|
| Title | Activity Levels Are Higher after Osteochondral Autograft Transfer Mosaicplasty than after Microfracture for Articular Cartilage Defects of the Knee |
| Type of study | Case series (retrospective) |
| Quality of study | Only MF data extracted Good |
| No. of patients | 48 with full-depth lesions Analysed at 1, 2, 3 and 5 years' FU; mean FU 4.4 years (range 2–10 years) |
| Population | Age at MF, mean 32.5 years (range 15–46 years) Male/female: 32:16 Lesion mean size 2.55 cm ² (range 1.00–6.25) cm ² BMI 25.5 kg/m ² (range 21–31 kg/m ²) Defect locations <ul style="list-style-type: none"> ● Medial femoral condyle, <i>n</i> = 27 ● Lateral femoral condyle, <i>n</i> = 16 ● Trochlea, <i>n</i> = 5 |
| Intervention | MF |
| Duration of injury? | NR |
| Previous attempts at repair? | None |
| Size of defect in cm ² | Mean 2.55 cm ² (range 1.00 to 6.25 cm ²) |
| Depth or severity if given | Full-depth lesions |
| Duration of FU | Mean FU 4.4 years (range 2–10 years) |
| Survival curve provided? | No |

| Krych 2012 ¹¹² | Data |
|--|---------------------|
| Results | |
| Definitions of success and failure | NR |
| SF-36 physical component mean (SD)SD 10 read from graph | Preop., 40.5 (10) |
| | Year 1, 47.9 (10) |
| | Year 2, 50.8 (10) |
| | Year 3, 52.6 (10) |
| | Year 5, 52.0 (10) |
| The Knee Outcome Survey ADL score mean (SD)SD read from graph | Preop., 64.1 (16) |
| | Year 1, 78.7 (19) |
| | Year 2, 79.1 (16) |
| | Year 3, 86.6 (13.4) |
| | Year 5, 84.4 (15.6) |
| SD read from graph | Preop., 49.7 (16) |
| | Year 1, 65.4 (16) |
| | Year 2, 69.2 (24) |
| | Year 3, 69.2 (25) |
| | Year 5, 84.4 (26) |
| Marx Activity Rating Scale score, mean (SD) | Preop., 7.3 (5.4) |
| | Year 1, 4.11 (1.05) |
| | Year 2, 3.71 (1.64) |
| | Year 3, 2.91 (2.12) |
| | Year 5, 2.89 (2.5) |
| Subgroup data: none reported | |
| Losses to FU: percentage and reasons if given | NR |
| Any costs given? | No |
| Survival curve? | No |
| FU, follow-up; NR, not reported; preop., preoperative. | |

Quality Assessment Tool for Case Series Studies: NIH

| Criteria | Yes | No | Other (CD, NR, NA) |
|---|-----|----|--------------------|
| 1. Was the study question or objective clearly stated? | ✓ | | |
| 2. Was the study population clearly and fully described, including a case definition? | ✓ | | |
| 3. Were the cases consecutive? | | | CD |
| 4. Were the subjects comparable? | | | NA |
| 5. Was the intervention clearly described? | ✓ | | |
| 6. Were the outcome measures clearly defined, valid, reliable and implemented consistently across all study participants? | ✓ | | |
| 7. Was the length of FU adequate? | ✓ | | |
| 8. Were the statistical methods well described? | ✓ | | |
| 9. Were the results well described? | ✓ | | |

QUALITY RATING: GOOD

CD, cannot determine; FU, follow-up; NA, not applicable; NR, not reported.

Moseley 2010

| Moseley 2010 ¹¹⁵ | Data |
|-----------------------------|--|
| Title | Long-Term Durability of Autologous Chondrocyte Implantation: A Multicenter, Observational Study in US Patients |
| Type of study | Case series |
| Quality of study | Fair to good |
| No. of patients | 72 |
| Population | <p>$N = 72$</p> <p>Mean FU (years), 10.9, SD 1.1</p> <p>Mean age (years) 37.0 ± 9.27, range 14–53</p> <p>Male (%) 61</p> <p>% with single defect, 60/72</p> <p>% with multiple defects, 12/72</p> <p>BMI mean \pm SD, 27.2 kg/m² range 13.2–42.4 kg/m²</p> <p>Defect size:</p> <ul style="list-style-type: none"> Total surface area, cm²: mean 5.2, range 0.4–23.5 <p>Defect sites (total defects = 84):</p> <ul style="list-style-type: none"> Medial femoral %, 72 Lateral femoral %, 18 Trochlea %, 10 |
| Intervention | <p>Carticel (Genzyme) ACP</p> <p>ACP received on or before 1996; 2044/2194 excluded because ACI treatment occurred after 31 December 1996</p> |

| Moseley 2010 ¹¹⁵ | Data |
|--|--|
| Duration of injury? | NR; 47/62 had acute onset of injury |
| Previous attempts at repair? | Previous intervention (in previous 5 years), %: |
| (Do not count debridement and lavage – only previous MF, abrasion, drilling, ACI) | <ul style="list-style-type: none"> At least 1 surgical procedure, 74% At least 1 cartilage repair procedure, 68% Abrasion/drilling/MF (MST), 36% Meniscus repair/meniscectomy, 28% |
| Size of defect, cm ² | 5.2, range 0.4–23.5 |
| Depth or severity if given | Full-thickness defects |
| Duration of FU | 6–10 years |
| Survival curve provided? | Yes |
| Results | |
| Failure | <p>Failure defined as: 'patient needed an operation after ACI that necessitated removal of the graft, confirmed a loss of defect fill, or violated the subchondral bone (e.g. abrasion chondroplasty, MF, drilling, unicompartmental knee replacement, total knee'</p> <p>Failures = 12/72</p> <p>18 patients who did not meet the definition of failure had operations for: presence of fibrotic tissue (4), periosteal flap complications (4), graft hypertrophy (3), adhesions (3), loose body (2), synovitis (2) and maltracking (2)</p> |
| OCS: A 1–10 VAS, with status allocated to scores of 2, 4, 6, 8, 10 defined, respectively as follows: | <ul style="list-style-type: none"> Improved at 1–5 years and at 6–10 years, $n = 47$ Improved at 1–5 years and not at 6–10 years, $n = 7$ Not improved at 1–5 years improved at 6–10 years, $n = 3$ Not improved at 1–5 years or at 6–10 years, $n = 15$ |
| <ul style="list-style-type: none"> Poor: I have significant limitations that affect ADL Good: I have moderate limitations that affect ADL, no sports possible Very good: I have only a few limitations with sports Excellent: I am able to do whatever I wish (any sport) with no problems | <p>No improvement from baseline was defined as a negative change or no change in OCS from baseline to latest FU. Improvement was defined as a positive score change of at least 1 point from baseline to latest FU</p> |
| Pain (mean SD) 1–10 VAS | <ul style="list-style-type: none"> Preop., 3.3 (3), $n = 72$ Year 1–5, 6.1 (3), $n = 72$ Year 6–10, 5.3 (3), $n = 72$ <p>Improved patients only:</p> <ul style="list-style-type: none"> Year 1–5, 7.5 (2), $n = 50$ Year 6–10, 7.4 (2.5), $n = 39$ |
| Swelling (mean SD) 1–10 VAS | <ul style="list-style-type: none"> Preop., 4.3 (3), $n = 72$ Year 1–5, 6.8 (4.4), $n = 72$ Year 6–10, 6.0 (4.5), $n = 72$ <p>Improved patients only:</p> <ul style="list-style-type: none"> Year 1–5, 7.5 (2), $n = 50$ Year 6–10, 7.4 (2.5), $n = 39$ |
| Subgroup data given? | Satisfaction according to defect site subgroups |
| Losses to FU: percentage and reasons if given | These covariate analyses were likely to be underpowered |
| Any costs given? | No |

| Moseley 2010 ¹¹⁵ | Data |
|--|--------------------------------------|
| Only for papers with survival curves | |
| Is curve KM? | Yes |
| If not, what is it? | |
| Risk table attached? | No |
| Total events reported? | Yes |
| HRs, <i>p</i> -value and/or 95% CI, and whether adjusted or not | NA, no subgroups analysed, so no HRs |
| FU, follow-up; NA, not applicable; NR, not reported; OCS, overall condition score; preop., preoperative. | |

Quality Assessment Tool for Case Series Studies: NIH

| Criteria | Yes | No | Other (CD, NR, NA) |
|---|-----|----|--------------------|
| 1. Was the study question or objective clearly stated? | ✓ | | |
| 2. Was the study population clearly and fully described, including a case definition? | ✓ | | |
| 3. Were the cases consecutive? | | ✓ | |
| 4. Were the subjects comparable? | | | NA |
| 5. Was the intervention clearly described? | ✓ | | |
| 6. Were the outcome measures clearly defined, valid, reliable and implemented consistently across all study participants? | | | CD |
| 7. Was the length of FU adequate? | ✓ | | |
| 8. Were the statistical methods well described? | ✓ | | |
| 9. Were the results well described? | ✓ | | |

QUALITY RATING: FAIR TO GOOD

CD, cannot determine; FU, follow-up; NA, not applicable; NR, not reported.

Additional comments

The large number of losses to FU is worrying.

Nawaz 2014

| Nawaz 2014 ⁸⁰ | Data |
|--------------------------|--|
| Title | Autologous Chondrocyte Implantation in the Knee: Mid-Term to Long-Term Results |
| Type of study | Case series |
| Quality of study | Good |
| No. of patients | 869 met inclusion criteria |
| | 41 lost to FU (one died before study) |
| | 827 analysed |
| Population | <i>N</i> = 827 |
| | Mean FU (years) 6.2 [2–12] |
| | Mean age (years) 34 [14–56] |
| | Male (%) 59.6 |

| Nawaz 2014 ⁸⁰ | Data |
|---|---|
| | Defect size (cm ²) 4.09 [0.64–20.7] |
| | Previous intervention 34% |
| | Defect site: <ul style="list-style-type: none"> • Medial femoral, 51% • Lateral femoral, 13% • Patellar, 24.0% • Trochlear, 6% • Multisite, 6% |
| Intervention | ACI-P/ACI-C/MACI |
| Duration of injury? | NR |
| Previous attempts at repair? | 34% not including debridement and lavage – only previous MF, abrasion, drilling, ACI |
| Size of defect in cm ² | Size, see above. Patients with defect with estimated depth of > 8 mm were not included. Lesions in the target population described as ‘regardless of depth or size’ |
| Depth or severity if given | |
| Duration of FU | See above |
| Survival curve provided? | Yes |
| Results | |
| Failure | Presented in KM plots Data extracted elsewhere |
| Stanmore functional rating (mean) | Preop. 2.7 |
| The <i>p</i> -value from ANOVA adjusted for time of postoperative estimate | Postop. 1.7 Mean difference –1.09 95% CI –1.18 to –1.00 <i>p</i> < 0.001 |
| VAS (0–10) | Preop. 5.95 |
| The <i>p</i> -value from ANOVA adjusted for time of postoperative estimate | Postop. 3.561 Mean difference –2.39 95% CI –2.61 to –2.19 <i>p</i> < 0.001 |
| Modified Cincinnati (0–100) <i>p</i> -value from ANOVA adjusted for time of postoperative estimate | Preop. 46.91 Postop. 66.74 Mean difference 19.83 95% CI 18.1 to 21.56 <i>p</i> < 0.001 |
| Complications | NR |
| Subgroup data given? | Yes for KM plots of failure |
| Losses to FU: percentage and reasons if given | 41 lost to FU, 1 died; 869 – 42 = 827 analysed |
| Any costs given? | No |

| Nawaz 2014 ⁸⁰ | Data |
|---|--|
| Only for papers with survival curves | |
| Is curve KM? | Yes. Several by subgroup |
| If not, what is it? | |
| Risk table attached? | To some |
| Total events reported? | For some |
| HRs, <i>p</i> -value and/or 95% CI, and whether adjusted or not | Yes for subgroup analyses; multivariate Cox regression |
| ANOVA, analysis of variance; FU, follow-up; NR, not reported; postop., postoperative; preop., preoperative. | |

Quality Assessment Tool for Case Series Studies: NIH

| Criteria | Yes | No | Other (CD, NR, NA) |
|---|-----|----|--------------------|
| 1. Was the study question or objective clearly stated? | ✓ | | |
| 2. Was the study population clearly and fully described, including a case definition? | ✓ | | |
| 3. Were the cases consecutive? | | ✓ | |
| 4. Were the subjects comparable? | ✓ | | |
| 5. Was the intervention clearly described? | ✓ | | |
| 6. Were the outcome measures clearly defined, valid, reliable and implemented consistently across all study participants? | ✓ | | |
| 7. Was the length of FU adequate? | ✓ | | |
| 8. Were the statistical methods well described? | ✓ | | |
| 9. Were the results well described? | ✓ | | |

QUALITY RATING: GOOD

CD, cannot determine; FU, follow-up; NA, not applicable; NR, not reported.

Niemeyer 2014¹¹⁷

| Niemeyer 2014 ¹¹⁷ | Data |
|------------------------------|---|
| Title | Long-term Outcomes After First-Generation Autologous Chondrocyte Implantation for Cartilage Defects of the Knee |
| Type of study | Case series |
| Quality of study | Good |
| No. of patients | 70 |
| Population | <i>N</i> = 70 |
| | 16 were lost to FU |
| | Mean FU (years) 10.9 (SD 1.1) |
| | Mean age (years) 33.3 (SD 10.2) |
| | Male (%) 35.7 |
| | Defect size (cm ²) 6.5 (SD 4.0) |
| | Previous intervention (%) 62.8 |

| Niemeier 2014 ¹¹⁷ | Data |
|---|--|
| | Defect site: <ul style="list-style-type: none"> • Medial femoral, 41.1% • Lateral femoral, 18.6% • Patellar, 20% • Trochlear, 2.9% • Multisite, 17.1% |
| Intervention | First-generation ACP |
| Duration of injury? | 'the mean duration of symptoms was several years' |
| Previous attempts at repair? | (44/70) 62.8% had previous intervention |
| | 20/44 were not defect associated |
| Size of defect in cm ² | Size, see above |
| Depth or severity if given | Full-thickness defects |
| | Defects of the subchondral bone plate exceeding a depth of 3–4 mm were excluded |
| Duration of FU | See above |
| Survival curve provided? | Yes |
| Results | |
| Failure | KM plot |
| VAS pain (mean SD) | At FU, pain at exposure on the VAS decreased from 7.2 ± 1.9 preop. to 2.1 ± 2.1 postop. ($p < 0.01$) |
| Lysholm (mean SD) | 42.0 ± 22.5 preop. to 71. ± 17.4 postop. |
| IKDC (mean SD) | FU 74.0 ± 17.3 |
| Tegner score (mean SD) | Decreased from 5.67 ± 2.39 to 4.36 ± 1.63 ($p < 0.01$) |
| | This represents slight worsening |
| FU scores: mean (SD) | |
| KOOS version 4 | 68.4 ± 19.9 |
| KOOS-pain | 81.4 ± 18.2 |
| KOOS-symptoms | 75.6 ± 17.3 |
| KOOS-ADL | 86.0 ± 16.7 |
| KOOS-sports | 62.3 ± 29.0 |
| KOOS-QoL | 54.3 ± 23.9 |
| Satisfaction (at FU), <i>n</i> | Very satisfied, 28 |
| | Satisfied, 26 |
| | Neutral, 14 |
| | Not satisfied, 2 |
| | Total, 70 |
| Complications | No complications related to the surgical procedure itself |
| Subgroup data given? | Satisfaction according to defect site subgroups |
| | Little difference, but numbers too small for conclusions |
| Losses to FU: percentage and reasons if given | 16 were lost to FU; no details |

| Niemeyer 2014 ¹¹⁷ | Data |
|--|---------------------------|
| Any costs given? | No |
| Only for papers with survival curves | |
| Is curve KM? | Yes |
| If not, what is it? | |
| Risk table attached? | No |
| Total events reported? | Yes |
| HRs, <i>p</i> -value and/or 95% CI, and whether adjusted or not | NA, no subgroups analysed |
| FU, follow-up; NA, not applicable; postop., postoperative; preop., preoperative. | |

Quality Assessment Tool for Case Series Studies: NIH

| Criteria | Yes | No | Other (CD, NR, NA) |
|---|-----|----|--------------------|
| 1. Was the study question or objective clearly stated? | ✓ | | |
| 2. Was the study population clearly and fully described, including a case definition? | ✓ | | |
| 3. Were the cases consecutive? | | ✓ | |
| 4. Were the subjects comparable? | ✓ | | |
| 5. Was the intervention clearly described? | ✓ | | |
| 6. Were the outcome measures clearly defined, valid, reliable and implemented consistently across all study participants? | ✓ | | |
| 7. Was the length of FU adequate? | ✓ | | |
| 8. Were the statistical methods well described? | ✓ | | |
| 9. Were the results well described? | ✓ | | |

QUALITY RATING: GOOD

CD, cannot determine; FU, follow-up; NA, not applicable; NR, not reported.

Additional comments

The large number of losses to FU is worrying.

Niemeyer 2014¹⁰²

| Niemeyer 2014 ¹⁰² | Data |
|------------------------------|--|
| Title | First-generation vs. Second-generation Autologous Chondrocyte Implantation for Treatment of Cartilage Defects of the Knee: A Matched-pair Analysis on Long-term Clinical Outcome |
| Type of study | Cohort with matched historical controls |
| Quality of study | Criteria for matching were defect location and patient age. If there were multiple options in the database, defect size was used an additional parameter for selection Good Five stars Newcastle/Ottawa |

| Niemeyer 2014 ¹⁰² | Data |
|--|--|
| No. of patients | N = 46 |
| Population | <ul style="list-style-type: none"> ● ACI-P = 23 – ACI-P were the historical controls ● ACI-C = 23 |
| Intervention | <p>Age (years) mean (SD):</p> <ul style="list-style-type: none"> ● ACI-P, 31.7 (6.9) ● ACI-C, 31.4 (7.8) <p>% male not reported</p> <p>Reason for injury not reported</p> <p>ACI-P (chondrocytes provided by Genzyme, Cambridge, MA, USA and Metreon Bioproducts GmbH, Freiburg, Germany)</p> <p>ACI-C (Chondro-Gide™, Geistlich, Wolhusen, Switzerland)</p> |
| Duration of injury? | NR |
| Previous attempts at repair? | NR |
| Size of defect in cm ² mean, (SD) | ACI-P: 5.1 (2.3) |
| Depth or severity if given | ACI-C: 4.9 (1.5) |
| Duration of FU, mean (SD) | All graded 3 or 4 according to the ICRS classification |
| | ACI-P: 10.7 (1.0) years |
| | ACI-C: 10.5 (0.6) years |
| Survival curve provided? | Yes |
| Results | |
| Re-intervention rate. Definition for re-intervention not given | <p>ACI-P: 4/23 (17.4%), including one total knee joint replacement</p> <p>ACI-C: 4/23 (17.4%) including one total knee joint replacement</p> |
| Lysholm score, mean SD | ACI-P: |
| | <ul style="list-style-type: none"> ● Preop., 38.4 (18.3) ● FU, 75.6 (11.8) |
| | ACI-C: |
| | <ul style="list-style-type: none"> ● Preop., 44.1 (21.3) ● FU, 82.7 (9.9) |
| | ACI-P vs. ACI-C preop.: $p = 0.371$ |
| | ACI-P vs. ACI-C at FU: $p = 0.031$ |
| No baseline data | ACI-P: 68.0 (12.0) ACI-C: 76.4 (12.8) $p = 0.023$ |
| Subgroup data given? | No subgroup data |
| Losses to FU: percentage and reasons if given | NA |
| Any costs given? | No |

| Niemeyer 2014 ¹⁰² | Data |
|---|------|
| Only for papers with survival curves | |
| Is curve KM? | Yes |
| If not, what is it? | |
| Risk table attached? | No |
| Total events reported? | No |
| HRs, <i>p</i> -value and/or 95% CI, and whether adjusted or not | No |
| FU, follow-up; NR, not reported; preop., preoperative. | |

Peterson 2010

| Peterson 2010 ¹¹⁸ | Data |
|-----------------------------------|---|
| Title | Autologous Chondrocyte Implantation: A Long-term Follow-up |
| Type of study | Case series. Retrospective data collection and analysis |
| Quality of study | Poor |
| No. of patients | 590 had ACI-P 341 eligible 224 responded to questionnaires Isolated cartilage lesions <i>n</i> = 159 Multiple lesions <i>n</i> = 56 |
| Population | Age 33.3 years (SD 9.5 years, range 14–61.5 years) % male not reported Reason for injury not reported |
| Intervention | ACI-P |
| Duration of injury? | NR |
| Previous attempts at repair? | 30/82 (37%) had a previous operation that included drilling or shaving of the chondral lesion Not clear what the <i>n</i> of 82 relates to |
| Size of defect in cm ² | 5.3 cm ² (range 0.6–15.8 cm ²) per lesion |
| Depth or severity if given | 7 cm ² (range 0.6–27 cm ²) per patient |
| Duration of FU | 12.8 years (range 9.3–20.7 years) |
| Survival curve provided? | No |
| Results | |
| | At FU: <ul style="list-style-type: none"> • Better or same: 165/224 (74%) • Worse: 59/224 (26%) Satisfied with ACI and would do again: 202/219 (92%) Success/failure not reported Current status during the past 10 years rated as better, worse, or unchanged (no further details) |

| Peterson 2010 ¹¹⁸ | Data |
|--|--|
| Lysholm score | Preop.: 60.3 FU: 69.5 ($p = 0.009$ from two-sample t -test; $p = 0.0016$ from paired t -test pertaining to 58 patients) |
| Tegner–Walgren score | Preop.: 7.22 FU: 8.2 ($p = 0.002$ from two-sample t -test; $p = 0.0008$ from paired t -test pertaining to 109 patients) |
| Brittberg–Peterson score | Preop.: 59.4 |
| 10 cm VAS with 13 parameters, where 0 relates to normal function and 130 severe disability | FU: 40.9 ($p < 0.001$ from two-sample t -test; $p = 0.004$ from paired t -test pertaining to 53 patients) |
| KOOS | No baseline data FU: Pain, 74.76 Symptoms, 63 ADL, 81 Sports, 41.5 QoL, 49.3 |
| Noyes score | FU: 5.4 |
| Subgroup data | |
| Improved compared with previous years, n (%) | Isolated femoral condyle defects ($n = 52$): 14 (27) Multiple lesions ($n = 55$): 12 (22) OCD lesions ($n = 26$): 7 (27) Patellar lesions with realignment ($n = 34$): 6 (18) Femoral condyle lesions with anterior cruciate ligament reconstruction ($n = 46$): 11 (24) |
| Same compared with previous years, n (%) | Isolated femoral condyle defects ($n = 52$): 22 (42) Multiple lesions ($n = 55$): 20 (40) OCD lesions ($n = 26$): 14 (54) Patellar lesions with realignment ($n = 34$): 18 (53) Femoral condyle lesions with anterior cruciate ligament reconstruction ($n = 46$): 24 (52) |
| Would do ACI again, n (%) | Isolated femoral condyle defects ($n = 52$): 47 (90) Multiple lesions ($n = 55$): 51 (94) OCD lesions ($n = 26$): 25 (96.2) Patellar lesions with realignment ($n = 34$): 31 (91.2) Femoral condyle lesions with anterior cruciate ligament reconstruction ($n = 46$): 41 (91.1) |

| Peterson 2010 ¹¹⁸ | Data |
|--|--|
| Lysholm score, mean (range) [available no. of values] | <p>Isolated femoral condyle defects ($n = 52$):</p> <ul style="list-style-type: none"> • Preop.: 60.1 (46–81) [13] • FU: 72.6 (25–96) • $p = 0.02$ (two-sample <i>t</i>-test) • $p = 0.03$ (paired <i>t</i>-test) <p>Multiple lesions ($n = 55$):</p> <ul style="list-style-type: none"> • Preop.: 50.9 [8] • FU: 67.7 (17–100) • $p = 0.05$ (two-sample <i>t</i>-test) • $p = 0.15$ (paired <i>t</i>-test) <p>OCD lesions ($n = 26$):</p> <ul style="list-style-type: none"> • Preop.: 56.2 (SD 22, range 13–85) [12] • FU: 67.4 (SD 16.4), (31–95) • $p = 0.1$ (two-sample <i>t</i>-test) • $p = 0.3$ (paired <i>t</i>-test) <p>Patellar lesions with realignment ($n = 34$):</p> <ul style="list-style-type: none"> • Preop.: 69 (47–85)⁶ • FU: 66 (17–100) • $p = 0.8$ (two-sample <i>t</i>-test) • $p = 0.3$ (paired <i>t</i>-test) <p>Femoral condyle lesions with anterior cruciate ligament reconstruction ($n = 46$):</p> <ul style="list-style-type: none"> • Preop.: 59.1 [16] • FU: 69.2 (34–100) • $p = 0.05$ (two-sample <i>t</i>-test) • $p = 0.1$ (paired <i>t</i>-test) |
| Tegner–Wallgren score, mean (range) [available no. of values] | <p>Isolated femoral condyle defects ($n = 52$):</p> <ul style="list-style-type: none"> • Preop.: 7.8 [2.6] • FU: 8 (2–14) • $p = 0.7$ (two-sample <i>t</i>-test) • $p = 0.7$ (paired <i>t</i>-test) <p>Multiple lesions ($n = 55$):</p> <ul style="list-style-type: none"> • Preop.: 7.2 [22] • FU: 8 (3–11) • $p = 0.1$ (two-sample <i>t</i>-test) • $p = 0.2$ (paired <i>t</i>-test) <p>OCD lesions ($n = 26$):</p> <ul style="list-style-type: none"> • Preop.: 6.4 (SD 2.2, range 1–9) [16] • FU: 8.6 (SD 1.6, range 5–13) • $p = 0.01$ (two-sample <i>t</i>-test) • $p = 0.03$ (paired <i>t</i>-test) <p>Patellar lesions with realignment ($n = 34$):</p> <ul style="list-style-type: none"> • Preop.: 7.4 (3–14) [17] • FU: 8.1 (3–14) • $p = 0.3$ (two-sample <i>t</i>-test) • $p = 0.2$ (paired <i>t</i>-test) |

Peterson 2010¹¹⁸

Data

| | | |
|--|--|--|
| | <p>Femoral condyle lesions with anterior cruciate ligament reconstruction ($n = 46$):</p> <ul style="list-style-type: none"> ● Preop.: 7.2 [33] ● FU: 8.1 (3–15) ● $p = 0.1$ (two-sample t-test) ● $p = 0.07$ (paired t-test) | |
| Brittberg–Peterson score, mean (range) [available no. of values] | <p>Isolated femoral condyle defects ($n = 52$):</p> <ul style="list-style-type: none"> ● Preop.: 65.9 (31–107) [12] ● FU: 38.4 (3–102.8) ● $p = 0.02$ (two-sample t-test) ● $p = 0.08$ (paired t-test) <p>Multiple lesions ($n = 55$):</p> <ul style="list-style-type: none"> ● Preop.: 64.1 [8] ● FU: 46.3 (1.7–115.8) ● $p = 0.12$ (two-sample t-test) ● $p = 0.9$ (paired t-test) <p>OCD lesions ($n = 26$):</p> <ul style="list-style-type: none"> ● Preop.: 51.8 (SD 32, range 9.4–104) [11] ● FU: 38.6 (SD 29, range 2.7–99) ● $p = 0.3$ (two-sample t-test) ● $p = 0.8$ (paired t-test) <p>Patellar lesions with realignment ($n = 34$):</p> <ul style="list-style-type: none"> ● Preop.: 50.1 (31–65) [6] ● FU: 49.2 (31–65) ● $p = 0.9$ (two-sample t-test) ● $p = 0.5$ (paired t-test) <p>Femoral condyle lesions with anterior cruciate ligament reconstruction ($n = 46$):</p> <ul style="list-style-type: none"> ● Preop.: 56.3 [14] ● FU: 41.1 (2–103.4) ● $p = 0.08$ (two-sample t-test) ● $p = 0.2$ (paired t-test) | |
| | <p>Isolated femoral condyle defects ($n = 52$):</p> <ul style="list-style-type: none"> ● Pain, 77.3 ● Symptoms, 65 ● ADL, 83.1 ● Sports, 45.1 ● QoL, 51 <p>Multiple lesions ($n = 55$):</p> <ul style="list-style-type: none"> ● Pain, 71.3 ● Symptoms, 61.5 ● ADL, 77.8 ● Sports, 37.4 ● QoL, 51 <p>OCD lesions ($n = 26$):</p> <ul style="list-style-type: none"> ● Pain, 78 ● Symptoms, 65.2 ● ADL, 85.6 ● Sports, 46.9 ● QoL, 54.3 | |
| | | |
| | KOOS | |

| Peterson 2010 ¹¹⁸ | Data |
|---|--|
| | Patellar lesions with realignment ($n = 34$): <ul style="list-style-type: none"> • Pain, 69.7 • Symptoms, 57.9 • ADL, 75 • Sports, 34.4 • QoL, 44.1 |
| | Femoral condyle lesions with anterior cruciate ligament reconstruction ($n = 46$): <ul style="list-style-type: none"> • Pain, 72.8 • Symptoms, 67.5 • ADL, 81.3 • Sports, 41.1 • QoL, 48.2 |
| Noyes score, mean range | Isolated femoral condyle defects ($n = 52$): 5.4 (1–9); states 5.4 in text, 5.3 in table Multiple lesions ($n = 55$): 5.2 (1–10) OCD lesions ($n = 26$): 5.7 (3–9) Patellar lesions with realignment ($n = 34$): 5.1 (1–10) Femoral condyle lesions with anterior cruciate ligament reconstruction ($n = 46$): 5.2 (1–9) |
| Losses to FU: percentage and reasons if given | 224/341 (65%) responded to questionnaires Only responders included in analysis |
| Any costs given? | No |
| Survival curve | No |

FU, follow-up; NR, not reported; preop., preoperative.

Quality Assessment Tool for Case Series Studies

| Criteria | Yes | No | Other (CD, NR, NA) |
|---|-----|----|--------------------|
| 1. Was the study question or objective clearly stated? | ✓ | | |
| 2. Was the study population clearly and fully described, including a case definition? | | ✓ | |
| 3. Were the cases consecutive? | | | CD |
| 4. Were the subjects comparable? | ✓ | | |
| 5. Was the intervention clearly described? | ✓ | | |
| 6. Were the outcome measures clearly defined, valid, reliable and implemented consistently across all study participants? | | ✓ | |
| 7. Was the length of FU adequate? | ✓ | | |
| 8. Were the statistical methods well described? | ✓ | | |
| 9. Were the results well described? | | ✓ | |

QUALITY RATING: FAIR

CD, cannot determine; FU, follow-up; NA, not applicable; NR, not reported.

Additional comments

Preoperative values not available for some outcomes. Baseline measures collected retrospectively from medical files.

Salzman 2013

| Salzman 2013 ¹³⁷ | Data |
|---|---|
| Title | Reoperative characteristics after MF of knee cartilage lesions in 454 patients |
| Type of study | Case series |
| Quality of study | Fair |
| No. of patients | 560 consecutive patients of which 454 were evaluated and 123 found to have been reoperated on the index lesion |
| Population | Mean FU for the 123 receiving reoperation was 5 years (SD 2.1) $n = 123$ Age at surgery: 44.2 ± 13.9 years Male/female: 67/56 BMI: 25.8 ± 3.6 kg/m ² Smoking/non-smoking: 30/93 |
| Intervention | MF |
| Duration of injury? | Symptom duration: 61.3 ± 68.6 months |
| Previous attempts at repair? | No. of previous surgeries: 1.9 ± 2.1 |
| Size of defect | # of defects, cm ² : 1.2 ± 0.5 |
| Depth or severity if given | Defect size/knee, cm ² : 2.1 ± 1.7 |
| All 123 had one defect or more; 22 had two defects; 2 had three defects | Defect #1 ($n = 123$), #2 ($n = 22$), #3 ($n = 2$) |
| So: | 99 with 1 defect, depth according to ICRS: |
| <ul style="list-style-type: none"> ● 99 had only one defect ● 22 had two defects ● 2 had three defects | <ul style="list-style-type: none"> ● ICRS °2# 1 ● ICRS °3B #26 ● ICRS °3C #36 ● ICRS °4 #36 |
| Commonest depth for defect: ICRS °3C and ICRS °3B | 22 with second defect (depth of largest): |
| Very few ICRS °2 or ICRS °4 | <ul style="list-style-type: none"> ● ICRS °2 #1 ● ICRS °3B #5 ● ICRS °3C #12 ● ICRS °4 #4 |
| | 2 with third defect (depth of largest): |
| | <ul style="list-style-type: none"> ● ICRS °2 #0 ● ICRS °3B #0 ● ICRS °3C #1 ● ICRS °4 #1 |
| Duration of FU | Mean FU for the 123 receiving reoperation was 5 years (SD 2.1) On average, reoperation commenced 18 months after initial MF |
| Survival curve provided? | No |

| Salzman 2013 ¹³⁷ | Data |
|---|---|
| Results | |
| Definitions of success and failure | Failure defined as above |
| Lysholm score mean (SD) | Preop. not reported Postop. Lysholm: 62.8 ± 24.5 |
| VAS knee pain, mean (SD) | Preop. NAS-P 3.1 ± 2.1, n = 123 Postop. NAS-P 5.2 ± 2.4, n = 123 |
| Numeric analogue scale (NAS) for pain (NAS-P), with 10 representing 'no pain' and 0 representing 'maximal imaginable pain.' | |
| VAS knee function, mean (SD) | NAS-F definition unclear <ul style="list-style-type: none"> • Preop. NAS-F 2.8 ± 1.8 • Postop. NAS-F 4.8 ± 2.2 |
| Subgroup data | |
| Failure | Findings based on regression analysis Failure was associated with the following factors: smaller lesions; more previous surgery; preoperative subjective sensation of less pain and less function; smoking; patellofemoral defects |
| VAS knee pain, mean (SD) | NR |
| VAS knee function, mean (SD) | NR |
| Losses to FU: percentage and reasons if given | Telephone interviews of some of the 560 patients were incomplete, leaving 454 for analysis |
| Any costs given? | No |
| Survival curve? | No |
| FU, follow-up; NAS, numeric analogue scale; NAS-F, numeric analogue scale for function; NAS-P, numeric analogue scale for pain; NR, not reported; postop., postoperative; preop., preoperative. | |

Quality Assessment Tool for Case Series Studies: NIH

| Criteria | Yes | No | Other (CD, NR, NA) |
|---|-----|----|--------------------|
| 1. Was the study question or objective clearly stated? | ✓ | | |
| 2. Was the study population clearly and fully described, including a case definition? | ✓ | | |
| 3. Were the cases consecutive? | | | CD |
| 4. Were the subjects comparable? | | | NA |
| 5. Was the intervention clearly described? | ✓ | | |
| 6. Were the outcome measures clearly defined, valid, reliable and implemented consistently across all study participants? | | ✓ | |
| 7. Was the length of FU adequate? | ✓ | | |
| 8. Were the statistical methods well described? | ✓ | | |
| 9. Were the results well described? | ✓ | | |

QUALITY RATING: FAIR

CD, cannot determine; FU, follow-up; NA, not applicable; NR, not reported.

Shive 2015

| Shive 2015 ¹²² | Data | | |
|--|--|------------------|------------------|
| Title | BST-CarGel® (Piramal Life Sciences, Mumbai, India) Treatment Maintains Cartilage Repair Superiority over Microfracture at 5 Years in a Multicenter Randomised Controlled Trial | | |
| Type of study | RCT | | |
| Quality of study | Fair | | |
| No. of patients | 80 originally randomised; this report $n = 60$ | | |
| Population: | | | |
| The trial randomised 41 and 39 to BST-CarGel and MF, respectively; these data are only for those followed to 5 (?) years | <i>N</i> | 34 | 26 |
| | Mean FU | NR | NR |
| Inclusion if single, focal cartilage lesion on the femoral condyles and moderate knee pain (> 4 on a 10-cm VAS) | Age mean (SD), years | 34.3 (9.7) | 40.1 (10.1) |
| | Male, % | 64.7 | 53.8 |
| Defect size, cm ² | mean (SD) | 2.41 (1.5) | 2.08 (1.22) |
| | maximum | 6.77 | 4.46 |
| | BMI (kg/m ²) mean (SD) | 27.6 (2.7) | 25.7 (2.9) |
| Symptom duration years | Median [range] | 1.4 [0.1–19.6] | 3 [0.3–27.8] |
| | Activity level <i>n</i> (%) | | |
| | High | 16 (47.1) | 15 (57.5) |
| | Medium | 16 (47.1) | 11 (42.3) |
| | Low | 2 (5.8) | 0 (0) |
| Previous intervention | NR | NR | |
| Intervention | MF or enhanced MF with BST-CarGel | | |
| | Multiple surgeons | | |
| Duration of injury? | See above | | |
| Previous attempts at repair? | NR | | |
| Size of defect in cm ² | See above | | |
| Depth or severity if given | Full thickness | | |
| Duration of FU | Appears to be 5 years | | |
| Survival curve provided? | No | | |
| Results | | | |
| Failure | NR | | |
| Lesion% fill | | BST-CarGel | MF |
| Least squares means \pm SE | <i>n</i> | 34 | 26 |
| | % fill | 93.79 \pm 1.16 | 86.96 \pm 2.85 |
| | $p = 0.017$. | | |

| Shive 2015 ¹²² | | Data | |
|---|---|-------------------|-------------------|
| WOMAC: | | BST-CarGel | MF |
| Change from baseline (least squares means \pm SE adjusted for baseline) | | | |
| | Pain | | |
| | <i>n</i> | 33 | 26 |
| | Score | -15.37 \pm 1.47 | -16.56 \pm 1.19 |
| | Stiffness | | |
| | <i>n</i> | 33 | 26 |
| | Score | -5.63 \pm 0.72 | -6.68 \pm 0.58 |
| | Physical function | | |
| | <i>n</i> | 33 | 26 |
| | Score | -56.52 \pm 4.57 | -62.10 \pm 3.43 |
| | No significant differences between groups. | | |
| SF-36: | Physical component: | | |
| Change from baseline (least squares means \pm SE adjusted for baseline) | | BST-CarGel | MF |
| | <i>n</i> | 34 | 27 |
| | Score | 13.12 \pm 1.63 | 14.48 \pm 1.42 |
| | Mental component: | | |
| | | BST-CarGel | MF |
| | <i>n</i> | 34 | 27 |
| | Score | 2.72 \pm 1.30 | -0.17 \pm 1.76 |
| | No significant differences between groups. | | |
| Mean T2 MRI relaxation time (ms) (least squares means \pm SE) | | BST-CarGel | MF |
| | <i>n</i> | 29 | 22 |
| | Score | 75.68 \pm 5.25 | 90.41 \pm 6.56 |
| | Aberrant data points for some patients were discarded. $p = 0.026$. | | |
| Complications | NR | | |
| Subgroup data given? | No | | |
| Losses to FU: percentage and reasons if given | 25% of patients lost to FU | | |
| Any costs given? | No | | |
| Survival curve? | No | | |
| FU, follow-up; NR, not reported. | | | |

Shive 2015 Quality Assessment Tool for RCT

| Bias | Author judgement | Support for judgement |
|---|------------------|--|
| Random sequence generation (selection bias) | Low risk | Via telephone interactive voice response system with use of a central, computer-generated randomisation schedule |
| Allocation concealment (selection bias) | Unclear risk | NR |
| Blinding of participants and personnel (performance bias) | Low risk | MRI assessments were blinded |
| Blinding of outcome assessment (detection bias) | Unclear risk | Patients not blinded because of incision size, bias in responses to questionnaires possible |
| Incomplete outcome data (attrition bias) | High risk | 25% of patients missing at 5-year FU |
| Selective reporting (reporting bias) | Unclear risk | Some non-primary outcomes appear to have been selected |
| Other bias | Low risk | None identified |

FU, follow-up; NR, not reported.

Solheim 2014

| Solheim 2014 ¹²⁴ | Data |
|------------------------------|--|
| Title | Results at 10–14 years after microfracture treatment of articular cartilage defects in the knee Follow-up to 2010 paper |
| Type of study | Case series |
| Quality of study | Fair 12 years' median FU (range 10–14 years) was reported for 110 (?) patients. Because baseline values differ between 2010 and 2014 papers, it is possible fewer than 110 were analysed in 2014 |
| No. of patients | 2010 paper: 116 eligible, 110 included in analysis; median age 38 years (range 15–60) <i>2014 paper: Included patients aged 60 years or younger</i> Patients having had a knee replacement (in the ipsilateral knee during the observation period) were denoted as failure, and their outcome score was not included in the calculations (of Lysholm score and VAS outcomes) |
| Population | Age 38 years (range 15–60 years) 58% male Reason for injury not reported Based on 110 analysed |
| Intervention | MF |
| Duration of injury? | Median 40 months (range 1 month to 20 years) |
| Previous attempts at repair? | NR |

| Solheim 2014 ¹²⁴ | Data |
|--|--|
| Size of defect in cm ² | One ($n = 76$), two ($n = 27$) or three ($n = 7$) |
| Depth or severity if given | ($n = 7$) lesions with a median total treated area of 4 cm ² (range 1–15 cm ²) |
| | Subgroups: |
| | <ul style="list-style-type: none"> • Single cartilage defect: mean 3.8 cm² (SD 1.5) • Multiple defects mean 7.5 cm² (SD 3.0) |
| Duration of FU | Median 5 years (range 2–9 years) |
| | <i>Median 12 (range 10–14) in 2014 paper</i> |
| Survival curve provided? | No (lacking both publications) |
| Results | |
| Definitions of success and failure: | Failures: 24/110 (22%) |
| Failure defined as a new surgical procedure with the intention to treat the cartilage lesion Lysholm score (e.g. another cartilage repair procedure, an osteotomy or a knee replacement) | Improved Lysholm score in non-failures: 67/86 (78%). Definition of 'improved' not reported |
| | <i>The 2014 paper: Patients having had a knee replacement (in the ipsilateral knee during the observation period) were denoted as failure, and their outcome score was not included in the calculations (of Lysholm score and VAS outcomes)</i> |
| | <i>The percentage patients with Failure/poor result was 47% (at medium term) and 45.5% (at 10–14 years). Failure ($n = 7$) was defined as above, and 'poor result' was defined as a Lysholm score of 64 or less or having a knee replacement</i> |
| Lysholm score mean (SD) | Preop.: 51 (18) |
| | FU: 71 (23); $p < 0.001$ |
| | <i>In 2014 paper:</i> |
| | <ul style="list-style-type: none"> • Preop.: 49 (18) • 'Medium FU' in 2014, 67 (23) • FU (10–14 years) 65 (24) The number of failures = 7 (omitted from calculation) • 'Medium FU' in 2014, 67 (23) [not 71 (23) as in 2010] • Presumably: 65 – 7 = 58 'poor' but without knee replacement at 10–14 years' FU |
| VAS knee pain, mean (SD) | Preop.: 52 (22) |
| Grading of knee pain and function of the knee | FU: 30 (24); $p < 0.001$ |
| | <i>In 2014 paper:</i> |
| By patient-administered VASs (VAS): 0 = no pain to 100 = worst possible pain | <ul style="list-style-type: none"> • Preop.: 55(21) • Medium FU 34 (24) • FU (10–14 years) 31 (24) The number of failures = 7 (omitted from calculation) • 2014 medium FU 34 (24) not same as 2010 [30 (24)] |
| VAS knee function, mean (SD) | Preop.: 41 (23) |
| VAS function: | FU: 69 (22); $p < 0.001$ |
| 0 = useless to 100 = full function | <i>In 2014 paper:</i> |
| | <ul style="list-style-type: none"> • Preop.: 40 (22) • Medium FU 63 (23) • FU (10–14 years) 65 (28) The number of failures = 7 (omitted from calculation) |
| Subgroup data | Failures: |
| | <ul style="list-style-type: none"> • Single chondral lesion 14/76 (18%) • Multiple lesions 10/34 (29%) |

| Solheim 2014 ¹²⁴ | Data |
|---|---|
| <p>Lysholm score mean (SD)</p> | <p>Improved Lysholm score in non-failures:</p> <ul style="list-style-type: none"> ● Single chondral lesion, 50/62 (18%) ● Multiple lesions, 17/24 (29%) <p>Single defect:</p> <ul style="list-style-type: none"> ● Preop.: 53 (17) ● FU: 74 (21); $p < 0.001$ <p>Multiple defects:</p> <ul style="list-style-type: none"> ● Preop.: 46 (21) ● FU: 63 (24); $p = 0.005$ <p>In 2014 paper:</p> <ul style="list-style-type: none"> ● A subgroup ($n = 30$) with ≤ 40 preoperative score had a poorer 10–14 years score than > 40 preoperative group: 56 (24) vs. 68 (22); $p = 0.02$ ● No relationship found between preoperative age or size of defect and 10–14 years FU score ● Poor outcome (score of ≤ 64) at 10–14 years (in 50 of 110, 45.5%) was associated with following subgroups: (a) signs of degenerative change around lesion at time of surgery (signs 54% poor vs. no signs 34% poor, $p = 0.04$); (b) previous or concurrent partial medial meniscectomy in ipsilateral knee (59% poor vs. 40% poor; $p = 0.048$); (c) $a \leq 40$ preoperative Lysholm score (60% poor vs. 39%; $p = 0.047$); (d) ≥ 36 months' preoperative duration of symptoms (52% poor vs. 30%; $p = 0.047$) ● Percentages only reported (no n/N data) ● If 110 were analysed the N in each subgroup would appear to be: <ul style="list-style-type: none"> ● (a) Signs = 63, no signs = 47 ● (b) Meniscectomy = 32%, no meniscectomy = 78% ● (c) $\leq 40 = 34^*$, $> 40 = 76$ ● (d) ≥ 36 months = 32, < 36 months = 78 ● *This number should be 30, the discrepancy may be due to rounding of percentages and that fewer than 110 were in fact analysed |
| <p>VAS knee pain, mean (SD)</p> | <p>Single defect:</p> <ul style="list-style-type: none"> ● Preop.: 52 (22) ● FU: 26 (21); $p < 0.001$ <p>Multiple defects:</p> <ul style="list-style-type: none"> ● Preop.: 53 (22) ● FU: 41 (27); $p = 0.018$ |
| <p>VAS knee function, mean (SD)</p> | <p>Single defect:</p> <ul style="list-style-type: none"> ● Preop.: 41 (24) ● FU: 74 (19); $p < 0.001$ <p>Multiple defects:</p> <ul style="list-style-type: none"> ● Preop.: 40 (19) ● FU: 54 (24); $p = 0.009$ |
| <p>Losses to FU: percentage and reasons if given</p> | <p>6/116 (5.2%) excluded from analysis (two died, four lost to FU or refused)</p> |
| <p>Any costs given?</p> | <p>No</p> |
| <p>Survival curve?</p> | <p>No</p> |
| <p>FU, follow-up; NR, not reported; preop., preoperative.</p> | |

Quality Assessment Tool for Case Series Studies: NIH

| Criteria | Yes | No | Other (CD, NR, NA) |
|---|-----|----|--------------------|
| 1. Was the study question or objective clearly stated? | ✓ | | |
| 2. Was the study population clearly and fully described, including a case definition? | ✓ | | |
| 3. Were the cases consecutive? | | | CD |
| 4. Were the subjects comparable? | ✓ | | |
| 5. Was the intervention clearly described? | ✓ | | |
| 6. Were the outcome measures clearly defined, valid, reliable and implemented consistently across all study participants? | | ✓ | |
| 7. Was the length of FU adequate? | ✓ | | |
| 8. Were the statistical methods well described? | ✓ | | |
| 9. Were the results well described? | ✓ | | |

QUALITY RATING: FAIR

CD, cannot determine; FU, follow-up; NA, not applicable; NR, not reported.

Additional comments

'Improvement' on Lysholm scale not defined.

Steadman 2003

| Steadman 2003 ¹²⁶ | Data |
|-----------------------------------|---|
| Title | Outcomes of Microfracture for Traumatic Chondral Defects of the Knee: Average 11-Year Follow-up |
| Type of study | Case series |
| Quality of study | Fair |
| No. of patients | 72 (75 knees) met inclusion criteria 68 (71 knees) included in results analysis |
| Population | Mean age (range): 30.4 years (13–35 years) 66.2% male Reason for injury? Either traumatic or degenerative Acute 15 knees, chronic 56 knees |
| Intervention | MF |
| Duration of injury? | Mean 3.2 years (range 0.02–16.1 years) |
| Previous attempts at repair? | Unclear |
| Size of defect in cm ² | 2.77 cm ² (range 0.2–10 cm ²) |
| Depth or severity if given | Full thickness |
| Duration of FU | Mean 11.3 years (range 7–17 years) |
| Survival curve provided? | No |

| Steadman 2003 ¹²⁶ | | Data | |
|---|---|----------------|--------------|
| Results | | | |
| MF failure | Extremely low rate; definition of failure not clear | | |
| Questionnaire scales vary: final vs. preop. | All scores represent clinical improvement | | |
| | Mean | SD | Range |
| • Satisfaction 1–10 | 8.3 | 1.6 | 4 to 10 |
| • Pain 1–4 | –1.5 | 0.9 | –3 to 1 |
| • Swelling 1–4 | –1.5 | 1 | –3 to 1 |
| • ADL 1–10 | 2.8 | 2.6 | –3 to 8 |
| • Strenuous work 1–10 | 2.7 | 3 | –4 to 9 |
| • Sport 1–10 | 2.9 | 3.4 | –4 to 8 |
| | Mean | SD | Range |
| Tegner final vs. preop. 1–10 best | 2.7 | 1.7 | –1 to 6 |
| | Assume this is mean of the individual score changes | | |
| | Mean | SD | Range |
| Lysholm final vs. preop. 1–100 best | 30.1 | 12.3 | 4 to 61 |
| | Assume this is mean of the individual score changes | | |
| Satisfaction | See above | | |
| Complications | ‘No perioperative complications were related to the surgical procedure’ Others not reported | | |
| Subgroup data given? | | | |
| Lysholm | Multivariate linear regression | | |
| | Coefficient | p-value | |
| Age | –0.299 | 0.011 | |
| Chronicity | –0.084 | 0.466 | |
| Location | –0.226 | 0.066 | |
| Size of lesions | –0.146 | 0.225 | |
| | Age is only influential factor and has negative effect on Lysholm score | | |
| Losses to FU: percentage and reasons if given | Two reasons given; also two patients who were considered failures were not included in analyses | | |
| Any costs given? | No | | |
| Survival curve? | No | | |
| FU, follow-up; preop., preoperative. | | | |

Quality Assessment Tool for Case Series Studies: NIH

| Criteria | Yes | No | Other (CD, NR, NA) |
|---|-----|----|--------------------|
| 1. Was the study question or objective clearly stated? | ✓ | | |
| 2. Was the study population clearly and fully described, including a case definition? | ✓ | | |
| 3. Were the cases consecutive? | | ✓ | |
| 4. Were the subjects comparable? | ✓ | | |
| 5. Was the intervention clearly described? | ✓ | | |
| 6. Were the outcome measures clearly defined, valid, reliable and implemented consistently across all study participants? | ✓ | | |
| 7. Was the length of FU adequate? | ✓ | | |
| 8. Were the statistical methods well described? | ✓ | | |
| 9. Were the results well described? | ✓ | | |

QUALITY RATING: FAIR

CD, cannot determine; FU, follow-up; NA, not applicable; NR, not reported.

Additional comments

The selection of participants was clearly retrospective and not consecutive (i.e. 25% of 302 consecutive patients were included). There was no mention of any re-intervention after MF. Only two MFs were judged to be failures, but no clear criteria for failure were offered. Outcome measures were subjective and some designed for this study only (not validated). Some patients with poor outcome were omitted from analyses.

Vanlauwe 2011

| Vanlauwe 2011 ⁴³ | Data | | |
|------------------------------|--|-------------|--------------|
| Title | Five-Year Outcome of Characterised Chondrocyte Implantation versus Microfracture for Symptomatic Cartilage Defects of the Knee | | |
| Type of study | RCT | | |
| Quality of study | Uncertain risk of bias [based on risk of selection bias (Cochrane risk of bias tool); see below] | | |
| No. of patients | Total 112: ACI 57 (51 treated); MF 61 | | |
| Population | Note: CCI – six did not get treated | | |
| Duration of injury? | | CCI | MFR |
| Previous attempts at repair? | <i>n</i> | 57 | 61 |
| | Age, years | 33.9 ± 8.6 | 33.9 ± 8.5 |
| | Height, cm | 177.0 ± 8.5 | 176.5 ± 10.8 |
| | Weight, kg | 80.6 ± 13.3 | 78.3 ± 13.9 |
| | Male, <i>n</i> (%) | 41 (67) | 35 (61) |
| | Female, <i>n</i> (%) | 20 (33) | 22 (39) |
| | Duration since onset, years: median (range) | 1.57 (0–18) | 1.97 (0–18) |
| | Proportion with previous surgery | 77% | 88% |
| | No. (%) with previous surgeries = 0 | 14 (23) | 7 (12) |
| | No. (%) with previous surgeries = 1 | 34 (56) | 29 (51) |
| | No. (%) with previous surgeries ≥ 2 | 13 (21) | 21 (37) |
| | Defect size, cm ² | 2.4 ± 1.2 | 2.6 ± 1.0 |

| Vanlauwe 2011 ⁴³ | Data | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|--|---|--------------|------------------------|------------|---------------------|------------|--------------|--------------|--------------|-----------------------|-------|-----|--------------|--------------|-----------------------|-------|------|--------------|--------------|-----------------------|-------|------------------------|--------------|--------------|-----------------------|-------|-----|--------------|--------------|-----------------------|-------|--|--------------|--------------|------------------------|-------|
| Intervention | ACI-P: ChondroCelect MF: as Steadman | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Size of defect | See above, ICRS grade 3 or 4 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Depth or severity if given | Deep lesions | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Duration of FU | 5 years | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Survival curve provided? | Yes | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Results | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Failures | ACI: 7/51 (13.7%) MF 10/61 (16.4%) log-rank $p = 0.561$ Failure defined as re-intervention | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| KOOS | Change from baseline at 5 years | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | <table border="1"> <thead> <tr> <th></th> <th>ACI</th> <th>MF</th> <th>Difference (95% CI)</th> <th>p-value</th> </tr> </thead> <tbody> <tr> <td>Overall KOOS</td> <td>21.17 ± 2.88</td> <td>14.07 ± 2.54</td> <td>7.1 (-0.52 to 14.73)</td> <td>0.068</td> </tr> <tr> <td>ADL</td> <td>16.42 ± 2.97</td> <td>11.35 ± 2.62</td> <td>5.07 (-2.79 to 12.94)</td> <td>0.203</td> </tr> <tr> <td>Pain</td> <td>19.04 ± 3.17</td> <td>13.27 ± 2.74</td> <td>5.77 (-2.55 to 14.09)</td> <td>0.172</td> </tr> <tr> <td>Symptoms/ stiffness</td> <td>17.70 ± 2.82</td> <td>10.90 ± 2.52</td> <td>6.81 (-0.70 to 14.32)</td> <td>0.075</td> </tr> <tr> <td>QoL</td> <td>32.12 ± 4.30</td> <td>21.23 ± 3.87</td> <td>10.89(-0.59 to 22.38)</td> <td>0.062</td> </tr> <tr> <td>Function, sports and recreational</td> <td>32.50 ± 5.88</td> <td>22.98 ± 5.69</td> <td>9.52 (-6.87 to 25.90)</td> <td>0.25</td> </tr> </tbody> </table> | | ACI | MF | Difference (95% CI) | p -value | Overall KOOS | 21.17 ± 2.88 | 14.07 ± 2.54 | 7.1 (-0.52 to 14.73) | 0.068 | ADL | 16.42 ± 2.97 | 11.35 ± 2.62 | 5.07 (-2.79 to 12.94) | 0.203 | Pain | 19.04 ± 3.17 | 13.27 ± 2.74 | 5.77 (-2.55 to 14.09) | 0.172 | Symptoms/ stiffness | 17.70 ± 2.82 | 10.90 ± 2.52 | 6.81 (-0.70 to 14.32) | 0.075 | QoL | 32.12 ± 4.30 | 21.23 ± 3.87 | 10.89(-0.59 to 22.38) | 0.062 | Function, sports and recreational | 32.50 ± 5.88 | 22.98 ± 5.69 | 9.52 (-6.87 to 25.90) | 0.25 |
| | ACI | MF | Difference (95% CI) | p -value | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Overall KOOS | 21.17 ± 2.88 | 14.07 ± 2.54 | 7.1 (-0.52 to 14.73) | 0.068 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| ADL | 16.42 ± 2.97 | 11.35 ± 2.62 | 5.07 (-2.79 to 12.94) | 0.203 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Pain | 19.04 ± 3.17 | 13.27 ± 2.74 | 5.77 (-2.55 to 14.09) | 0.172 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Symptoms/ stiffness | 17.70 ± 2.82 | 10.90 ± 2.52 | 6.81 (-0.70 to 14.32) | 0.075 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| QoL | 32.12 ± 4.30 | 21.23 ± 3.87 | 10.89(-0.59 to 22.38) | 0.062 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Function, sports and recreational | 32.50 ± 5.88 | 22.98 ± 5.69 | 9.52 (-6.87 to 25.90) | 0.25 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| KOOS subgroup | Change from baseline at 5 years, patients with < 3 years of symptoms Pre-planned subgroup | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | <table border="1"> <thead> <tr> <th></th> <th>ACI</th> <th>MF</th> <th>Difference (95% CI)</th> <th>p-value</th> </tr> </thead> <tbody> <tr> <td>Overall KOOS</td> <td>25.96 ± 3.45</td> <td>15.28 ± 3.17</td> <td>10.69 (1.30 to 20.07)</td> <td>0.026</td> </tr> <tr> <td>ADL</td> <td>18.95 ± 3.46</td> <td>12.53 ± 3.18</td> <td>6.41 (-3.01 to 15.83)</td> <td>0.178</td> </tr> <tr> <td>Pain</td> <td>22.86 ± 3.66</td> <td>13.75 ± 3.30</td> <td>9.12 (-0.76 to 18.99)</td> <td>0.07</td> </tr> <tr> <td>Symptoms/ stiffness</td> <td>21.43 ± 3.47</td> <td>13.34 ± 3.19</td> <td>8.09 (-1.35 to 17.54)</td> <td>0.092</td> </tr> <tr> <td>QoL</td> <td>40.51 ± 5.47</td> <td>21.48 ± 5.03</td> <td>19.02 (4.14 to 33.91)</td> <td>0.013</td> </tr> <tr> <td>Function, sports and recreational activities</td> <td>40.15 ± 7.66</td> <td>24.85 ± 7.66</td> <td>15.29 (-6.65 to 37.23)</td> <td>0.166</td> </tr> </tbody> </table> | | ACI | MF | Difference (95% CI) | p -value | Overall KOOS | 25.96 ± 3.45 | 15.28 ± 3.17 | 10.69 (1.30 to 20.07) | 0.026 | ADL | 18.95 ± 3.46 | 12.53 ± 3.18 | 6.41 (-3.01 to 15.83) | 0.178 | Pain | 22.86 ± 3.66 | 13.75 ± 3.30 | 9.12 (-0.76 to 18.99) | 0.07 | Symptoms/ stiffness | 21.43 ± 3.47 | 13.34 ± 3.19 | 8.09 (-1.35 to 17.54) | 0.092 | QoL | 40.51 ± 5.47 | 21.48 ± 5.03 | 19.02 (4.14 to 33.91) | 0.013 | Function, sports and recreational activities | 40.15 ± 7.66 | 24.85 ± 7.66 | 15.29 (-6.65 to 37.23) | 0.166 |
| | ACI | MF | Difference (95% CI) | p -value | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Overall KOOS | 25.96 ± 3.45 | 15.28 ± 3.17 | 10.69 (1.30 to 20.07) | 0.026 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| ADL | 18.95 ± 3.46 | 12.53 ± 3.18 | 6.41 (-3.01 to 15.83) | 0.178 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Pain | 22.86 ± 3.66 | 13.75 ± 3.30 | 9.12 (-0.76 to 18.99) | 0.07 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Symptoms/ stiffness | 21.43 ± 3.47 | 13.34 ± 3.19 | 8.09 (-1.35 to 17.54) | 0.092 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| QoL | 40.51 ± 5.47 | 21.48 ± 5.03 | 19.02 (4.14 to 33.91) | 0.013 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Function, sports and recreational activities | 40.15 ± 7.66 | 24.85 ± 7.66 | 15.29 (-6.65 to 37.23) | 0.166 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| AEs | Over 5 years 42 (82%) and 38 (62%) ACI and MF patients experienced at least one TEAEs | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Mean Tegner score | NR | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Subgroup data | See above for KOOS | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| No. of failures by 5 years | See above | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Losses to FU | 6/57 in the ACI arm did not receive treatment | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Any costs given? | No | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

| Vanlauwe 2011 ⁴³ | Data |
|--|---|
| Only for papers with survival curves | |
| Is curve KM? | Yes |
| If not, what is it? | |
| Risk table attached? | Yes (but for the MF arm does not appear sensible) |
| Total events reported? | Yes |
| HRs, <i>p</i> -value and/or 95% CI, and whether adjusted | No |
| | Log-rank test <i>p</i> -value |
| | No adjustment |

FU, follow-up; NR, not reported; TEAE, treatment-emergent adverse event.

Cochrane risk of bias score

| Bias | Author judgement | Support for judgement |
|---|------------------|---|
| Random sequence generation (selection bias) | Unclear risk | Minimisation not fully described |
| Allocation concealment (selection bias) | Low risk | Allocation through an interactive voice response system |
| Blinding of participants and personnel (performance bias) | High risk | Patients not blinded bias likely but unclear |
| Blinding of outcome assessment (detection bias) | Unclear risk | MRI independent centre carrying out the analyses of primary end points was unaware of patient treatment |
| Incomplete outcome data (attrition bias) | Low risk | FU complete for treated patients |
| Selective reporting (reporting bias) | Low risk | KOOS and AEs prespecified and reported |
| Other bias | Unclear risk | Errors in risk table for KM plot |

FU, follow-up.

Appendix 9 Models of time to failure included published autologous chondrocyte implantation and microfracture studies

This appendix lists information criteria for models used in analyses of reconstructed KM plots and reconstructed IPD. Graphs-of-model fits for included studies that are most relevant to the decision problem are presented, arranged by study in alphabetical order. Other appendices provide model information for the MF study of Bae *et al.*,¹³⁴ Gudas *et al.*¹²⁸ and the unpublished ACTIVE trial.³⁵

Data from some studies were sparse and immature (a small proportion of participants experienced an event) and, using the specified methods, some models and/or model 95% CIs could not be computed. Cumulative hazard model tests are available from authors on request.

TABLE 56 The AIC and BIC information criteria for parametric models

| Bentley 2012⁷⁸ | | | | | |
|---|-------------|------------------|-----------|------------|------------|
| Model | Obs. | ll(model) | df | AIC | BIC |
| Gamma | 58 | -44.3728 | 3 | 94.74568 | 100.927 |
| Exponential | 58 | -46.9174 | 1 | 95.83471 | 97.89515 |
| Weibull | 58 | -46.7082 | 2 | 97.4164 | 101.5373 |
| Gompertz | 58 | -46.8558 | 2 | 97.71164 | 101.8325 |
| Log-normal | 58 | -45.6348 | 2 | 95.26963 | 99.39052 |
| Log-logistic | 58 | -46.3567 | 2 | 96.7133 | 100.8342 |
| Linearly increasing hazard (one parameter) | 58 | -49.498 | 1 | 100.996 | 103.0565 |
| Biant 2014⁷⁹ | | | | | |
| Model | Obs. | ll(model) | df | AIC | BIC |
| Gamma | 104 | -77.5841 | 3 | 161.1683 | 169.1014 |
| Exponential | 104 | -81.5033 | 1 | 165.0067 | 167.6511 |
| Weibull | 104 | -80.2895 | 2 | 164.579 | 169.8678 |
| Gompertz | 104 | -81.488 | 2 | 166.976 | 172.2648 |
| Log-normal | 104 | -78.3939 | 2 | 160.7879 | 166.0766 |
| Log-logistic | 104 | -79.5402 | 2 | 163.0805 | 168.3693 |
| Linearly increasing hazard (one parameter) | 104 | -82.9069 | 1 | 167.8137 | 170.4581 |
| Bathtub | 104 | -81.443 | 3 | 168.886 | 176.8191 |
| Linearly increasing hazard (two parameters) | 104 | -81.443 | 2 | 166.886 | 172.1748 |

continued

TABLE 56 The AIC and BIC information criteria for parametric models (continued)

| Knutsen 2007,⁶⁷ ACI | | | | | |
|---|-------------|------------------|-----------|------------|------------|
| Model | Obs. | ll(model) | df | AIC | BIC |
| Gamma | 40 | -28.8468 | 3 | 63.69364 | 68.76028 |
| Exponential | 40 | -29.6877 | 1 | 61.37538 | 63.06426 |
| Weibull | 40 | -29.5975 | 2 | 63.1949 | 66.57266 |
| Gompertz | 40 | -29.6668 | 2 | 63.33367 | 66.71143 |
| Log-normal | 40 | -29.1317 | 2 | 62.26343 | 65.64118 |
| Log-logistic | 40 | -29.4573 | 2 | 62.91451 | 66.29226 |
| Linearly increasing hazard (one parameter) | 40 | -31.4127 | 1 | 64.82539 | 66.51427 |
| Knutsen 2007,⁶⁷ MF | | | | | |
| Model | Obs. | ll(model) | df | AIC | BIC |
| Gamma | 40 | -25.7202 | 3 | 57.44034 | 62.50698 |
| Exponential | 40 | -27.0267 | 1 | 56.05329 | 57.74217 |
| Weibull | 40 | -25.7411 | 2 | 55.48211 | 58.85987 |
| Gompertz | 40 | -25.8565 | 2 | 55.71308 | 59.09084 |
| Log-normal | 40 | -25.7855 | 2 | 55.57099 | 58.94875 |
| Log-logistic | 40 | -25.7727 | 2 | 55.54536 | 58.92312 |
| Linearly increasing hazard (two parameters) | 40 | -25.801 | 2 | 55.60197 | 58.97973 |
| Linearly increasing hazard (one parameter) | 40 | -25.8169 | 1 | 53.63372 | 55.3226 |
| Layton 2015¹¹³ | | | | | |
| Model | Obs. | ll(model) | df | AIC | BIC |
| Flexible parametric | 3498 | -1988.591 | 2 | 3981.182 | 3993.502 |
| Gamma | 3498 | -1972.989 | 3 | 3951.978 | 3970.458 |
| Exponential | 3498 | -2080.367 | 1 | 4162.734 | 4168.894 |
| Weibull | 3498 | -1981.392 | 2 | 3966.784 | 3979.104 |
| Gompertz | 3498 | -1972.5 | 2 | 3948.999 | 3961.319 |
| Log-normal | 3498 | -1981.837 | 2 | 3967.675 | 3979.995 |
| Log-logistic | 3498 | -1988.591 | 2 | 3981.182 | 3993.502 |
| Linearly increasing hazard (one parameter) | 3498 | -1985.295 | 1 | 3972.59 | 3978.75 |
| Linearly increasing hazard (two parameters) | 3498 | -1985.276 | 2 | 3974.553 | 3986.873 |
| Minas 2014,¹⁶⁴ all | | | | | |
| Model | Obs. | ll(model) | df | AIC | BIC |
| Exponential | 210 | -196.368 | 1 | 394.7365 | 398.0836 |
| Weibull | 210 | -191.79 | 2 | 387.5798 | 394.274 |
| Gompertz | 210 | -185.425 | 2 | 374.8509 | 381.5451 |
| Log-normal | 210 | -187.581 | 2 | 379.162 | 385.8562 |
| Log-logistic | 210 | -190.557 | 2 | 385.1139 | 391.8081 |
| Linearly increasing hazard (one parameter) | 210 | -243.436 | 1 | 488.8725 | 492.2196 |

TABLE 56 The AIC and BIC information criteria for parametric models (continued)

| Minas 2014,¹⁶⁴ previous intervention | | | | | |
|---|-------------|------------------|-----------|------------|------------|
| Model | Obs. | ll(model) | df | AIC | BIC |
| Gamma | 89 | -92.1654 | 3 | 190.3308 | 197.7967 |
| Exponential | 89 | -100.147 | 1 | 202.2938 | 204.7824 |
| Weibull | 89 | -99.7299 | 2 | 203.4597 | 208.437 |
| Gompertz | 89 | -97.3564 | 2 | 198.7128 | 203.69 |
| Log-normal | 89 | -96.9522 | 2 | 197.9044 | 202.8817 |
| Log-logistic | 89 | -98.646 | 2 | 201.2921 | 206.2693 |
| Linearly increasing hazard (one parameter) | 89 | -118.007 | 1 | 238.0132 | 240.5019 |
| Minas 2014,¹⁶⁴ no-previous intervention | | | | | |
| Model | Obs. | ll(model) | df | AIC | BIC |
| Gamma | 121 | -87.4531 | 3 | 180.9063 | 189.2936 |
| Exponential | 121 | -94.5668 | 1 | 191.1335 | 193.9293 |
| Weibull | 121 | -93.769 | 2 | 191.5379 | 197.1295 |
| Gompertz | 121 | -91.2933 | 2 | 186.5866 | 192.1782 |
| Log-normal | 121 | -91.9782 | 2 | 187.9563 | 193.5479 |
| Log-logistic | 121 | -93.3413 | 2 | 190.6826 | 196.2742 |
| Linearly increasing hazard (one parameter) | 121 | -110.052 | 1 | 222.1033 | 224.8991 |
| Moseley 2010¹¹⁵ | | | | | |
| Model | Obs. | ll(model) | df | AIC | BIC |
| Gamma | 72 | -45.9957 | 3 | 97.99145 | 104.8214 |
| Exponential | 72 | -47.6374 | 1 | 97.27483 | 99.5515 |
| Weibull | 72 | -47.359 | 2 | 98.71793 | 103.2713 |
| Gompertz | 72 | -46.8926 | 2 | 97.78513 | 102.3385 |
| Log-normal | 72 | -46.8112 | 2 | 97.62248 | 102.1758 |
| Log-logistic | 72 | -47.2613 | 2 | 98.52252 | 103.0759 |
| Linearly increasing hazard (one parameter) | 72 | -54.408 | 1 | 110.8159 | 113.0926 |
| Nawaz 2014⁸⁰ | | | | | |
| Model | Obs. | ll(model) | df | AIC | BIC |
| Gamma | 827 | -568.507 | 3 | 1143.014 | 1157.167 |
| Exponential | 827 | -625.545 | 1 | 1253.089 | 1257.807 |
| Weibull | 827 | -570.836 | 2 | 1145.672 | 1155.108 |
| Gompertz | 827 | -586.411 | 2 | 1176.822 | 1186.258 |
| Log-normal | 827 | -569.915 | 2 | 1143.831 | 1153.266 |
| Log-logistic | 827 | -568.834 | 2 | 1141.668 | 1151.103 |
| Linearly increasing hazard (two parameters) | 827 | -571.708 | 2 | 1147.416 | 1156.851 |
| Linearly increasing hazard (one parameter) | 827 | -571.82 | 1 | 1145.64 | 1150.358 |

continued

TABLE 56 The AIC and BIC information criteria for parametric models (*continued*)

| Nawaz 2014,⁸⁰ previous intervention | | | | | |
|--|-------------|------------------|-----------|------------|------------|
| Model | Obs. | ll(model) | df | AIC | BIC |
| Gamma | 280 | -323.345 | 3 | 652.69 | 663.5944 |
| Exponential | 280 | -360.259 | 1 | 722.5175 | 726.1522 |
| Weibull | 280 | -335.845 | 2 | 675.6899 | 682.9595 |
| Gompertz | 280 | -351.191 | 2 | 706.3822 | 713.6518 |
| Log-normal | 280 | -323.529 | 2 | 651.058 | 658.3276 |
| Log-logistic | 280 | -323.9 | 2 | 651.7998 | 659.0694 |
| Linearly increasing hazard (one parameter) | 280 | -348.764 | 1 | 699.5281 | 703.1629 |
| Bathtub | 280 | -344.046 | 3 | 694.091 | 704.9954 |
| Linearly increasing hazard (two parameters) | 280 | -344.046 | 2 | 692.091 | 699.3606 |
| Nawaz 2014,⁸⁰ no previous intervention | | | | | |
| Model | Obs. | ll(model) | df | AIC | BIC |
| Gamma | 547 | -404.093 | 3 | 814.186 | 827.0994 |
| Exponential | 547 | -422.594 | 1 | 847.1882 | 851.4926 |
| Weibull | 547 | -413.556 | 2 | 831.1128 | 839.7217 |
| Gompertz | 547 | -421.083 | 2 | 846.1666 | 854.7755 |
| Log-normal | 547 | -405.991 | 2 | 815.9812 | 824.5901 |
| Log-logistic | 547 | -410.462 | 2 | 824.9231 | 833.532 |
| Bathtub | 547 | -418.881 | 3 | 843.7628 | 856.6761 |
| Linearly increasing hazard (two parameters) | 547 | -418.881 | 2 | 841.7628 | 850.3717 |
| Linearly increasing hazard (one parameter) | 547 | -423.034 | 1 | 848.0674 | 852.3718 |
| Nawaz 2014,⁸⁰ lateral femoral site | | | | | |
| Model | Obs. | ll(model) | df | AIC | BIC |
| Gamma | 109 | -80.7189 | 3 | 167.4378 | 175.5118 |
| Exponential | 109 | -84.7776 | 1 | 171.5553 | 174.2466 |
| Weibull | 109 | -83.4963 | 2 | 170.9926 | 176.3753 |
| Gompertz | 109 | -84.7494 | 2 | 173.4988 | 178.8815 |
| Log-normal | 109 | -81.5869 | 2 | 167.1738 | 172.5565 |
| Log-logistic | 109 | -82.759 | 2 | 169.5179 | 174.9006 |
| Linearly increasing hazard (one parameter) | 109 | -86.1576 | 1 | 174.3151 | 177.0065 |
| Linearly increasing hazard (two parameters) | 109 | -84.6778 | 2 | 173.3557 | 178.7384 |

TABLE 56 The AIC and BIC information criteria for parametric models (*continued*)

| Nawaz 2014,⁸⁰ medial femoral site | | | | | |
|---|-------------|------------------|-----------|------------|------------|
| Model | Obs. | ll(model) | df | AIC | BIC |
| Gamma | 421 | -466.36 | 3 | 938.719 | 950.8469 |
| Exponential | 421 | -489.691 | 1 | 981.3818 | 985.4245 |
| Weibull | 421 | -478.693 | 2 | 961.3867 | 969.4719 |
| Gompertz | 421 | -487.87 | 2 | 979.7396 | 987.8249 |
| Log-normal | 421 | -467.487 | 2 | 938.973 | 947.0583 |
| Log-logistic | 421 | -470.581 | 2 | 945.1612 | 953.2465 |
| Bathtub | 421 | -486.252 | 3 | 978.5033 | 990.6312 |
| Linearly increasing hazard (one parameter) | 421 | -501.922 | 1 | 1005.844 | 1009.887 |
| Linearly increasing hazard (two parameters) | 421 | -486.252 | 2 | 976.5033 | 984.5886 |
| Nawaz 2014,⁸⁰ multisite | | | | | |
| Model | Obs. | ll(model) | df | AIC | BIC |
| Gamma | 47 | -43.0205 | 3 | 92.04091 | 97.59135 |
| Exponential | 47 | -49.4276 | 1 | 100.8552 | 102.7053 |
| Weibull | 47 | -44.5392 | 2 | 93.07841 | 96.77871 |
| Gompertz | 47 | -46.4715 | 2 | 96.94301 | 100.6433 |
| Log-normal | 47 | -43.2675 | 2 | 90.5349 | 94.2352 |
| Log-logistic | 47 | -43.7604 | 2 | 91.52073 | 95.22102 |
| Linearly increasing hazard (one parameter) | 47 | -44.5599 | 1 | 91.11976 | 92.96991 |
| Linearly increasing hazard (two parameters) | 47 | -44.0869 | 2 | 92.17384 | 95.87413 |
| Nawaz 2014,⁸⁰ patella site | | | | | |
| Model | Obs. | ll(model) | df | AIC | BIC |
| Gamma | 200 | -213.659 | 3 | 433.3189 | 443.2138 |
| Exponential | 200 | -227.676 | 1 | 457.3519 | 460.6502 |
| Weibull | 200 | -216.182 | 2 | 436.3644 | 442.961 |
| Gompertz | 200 | -221.612 | 2 | 447.2244 | 453.821 |
| Log-normal | 200 | -213.703 | 2 | 431.4064 | 438.003 |
| Log-logistic | 200 | -213.828 | 2 | 431.6568 | 438.2534 |
| Linearly increasing hazard (two parameters) | 200 | -218.561 | 2 | 441.1217 | 447.7183 |
| Linearly increasing hazard (one parameter) | 200 | -220.296 | 1 | 442.5924 | 445.8907 |

continued

TABLE 56 The AIC and BIC information criteria for parametric models (*continued*)

| Nawaz 2014,⁸⁰ trochlea site | | | | | |
|---|-------------|------------------|-----------|------------|------------|
| Model | Obs. | ll(model) | df | AIC | BIC |
| Gamma | 50 | -46.1589 | 3 | 98.31787 | 104.0539 |
| Exponential | 50 | -48.5806 | 1 | 99.16116 | 101.0732 |
| Weibull | 50 | -46.3142 | 2 | 96.62834 | 100.4524 |
| Gompertz | 50 | -47.1127 | 2 | 98.22545 | 102.0495 |
| Log-normal | 50 | -46.2374 | 2 | 96.47485 | 100.2989 |
| Log-logistic | 50 | -46.1139 | 2 | 96.2278 | 100.0518 |
| Linearly increasing hazard (one parameter) | 50 | -46.8221 | 1 | 95.64415 | 97.55618 |
| Linearly increasing hazard (two parameters) | 50 | -46.5753 | 2 | 97.15064 | 100.9747 |
| Niemeyer 2014¹¹⁷ | | | | | |
| Model | Obs. | ll(model) | df | AIC | BIC |
| Gamma | 70 | -61.0638 | 3 | 128.1277 | 134.8732 |
| Exponential | 70 | -60.8876 | 1 | 123.7751 | 126.0236 |
| Weibull | 70 | -60.8874 | 2 | 125.7747 | 130.2717 |
| Gompertz | 70 | -60.8557 | 2 | 125.7113 | 130.2083 |
| Log-normal | 70 | -60.0888 | 2 | 124.1775 | 128.6745 |
| Log-logistic | 70 | -60.7415 | 2 | 125.4829 | 129.9799 |
| Linearly increasing hazard (one parameter) | 70 | -68.8165 | 1 | 139.6329 | 141.8814 |
| Linearly increasing hazard (two parameters) | 70 | -60.8622 | 2 | 125.7245 | 130.2215 |
| Vanlauwe 2011,⁴³ ACI | | | | | |
| Model | Obs. | ll(model) | df | AIC | BIC |
| Gamma | 51 | -21.6794 | 3 | 49.35883 | 55.1543 |
| Exponential | 51 | -23.3598 | 1 | 48.71968 | 50.65151 |
| Weibull | 51 | -21.681 | 2 | 47.3619 | 51.22555 |
| Gompertz | 51 | -21.8389 | 2 | 47.67779 | 51.54144 |
| Log-normal | 51 | -21.8216 | 2 | 47.6432 | 51.50685 |
| Log-logistic | 51 | -21.6851 | 2 | 47.3701 | 51.23375 |
| Linearly increasing hazard (one parameter) | 51 | -21.6856 | 1 | 45.37118 | 47.30301 |
| Linearly increasing hazard (two parameters) | 51 | -21.6531 | 2 | 47.30627 | 51.16992 |

TABLE 56 The AIC and BIC information criteria for parametric models (*continued*)

| Vanlauwe 2011,⁴³ MF | | | | | |
|---|-------------|------------------|-----------|------------|------------|
| Model | Obs. | ll(model) | df | AIC | BIC |
| Gamma | 61 | -32.4298 | 3 | 70.85961 | 77.19223 |
| Exponential | 61 | -35.7444 | 1 | 73.48888 | 75.59975 |
| Weibull | 61 | -35.628 | 2 | 75.25597 | 79.47772 |
| Gompertz | 61 | -35.6081 | 2 | 75.21612 | 79.43787 |
| Log-normal | 61 | -34.7726 | 2 | 73.54515 | 77.76689 |
| Log-logistic | 61 | -35.4313 | 2 | 74.86256 | 79.08431 |
| Linearly increasing hazard (one parameter) | 61 | -37.6677 | 1 | 77.3354 | 79.44627 |
| Linearly increasing hazard (two parameters) | 61 | -35.3329 | 2 | 74.66589 | 78.88763 |
| Saris 2009,⁷⁰ MF | | | | | |
| Model | Obs. | ll(model) | df | AIC | BIC |
| Gamma | 61 | -22.6833 | 3 | 51.36665 | 57.69927 |
| Exponential | 61 | -25.351 | 1 | 52.70196 | 54.81284 |
| Weibull | 61 | -23.947 | 2 | 51.89407 | 56.11582 |
| Gompertz | 61 | -24.5998 | 2 | 53.19964 | 57.42138 |
| Log-normal | 61 | -23.5357 | 2 | 51.07138 | 55.29313 |
| Log-logistic | 61 | -23.8799 | 2 | 51.75988 | 55.98163 |
| Linearly increasing hazard (one parameter) | 61 | -23.9471 | 1 | 49.89419 | 52.00507 |

AIC, Akaike information criterion; BIC, Bayesian information criterion; df, degrees of freedom; Obs., observations.

For Figures 49–74, unless stated otherwise, the following abbreviations apply: bt, bathtub; ex, exponential; ga, Gamma; go, Gompertz; ll, log-logistic; ln, log-normal; ord, ordinate; ra, two-parameter linearly increasing hazard model (Rayleigh); sq, single parameter linearly increasing hazard model; we, Weibull.

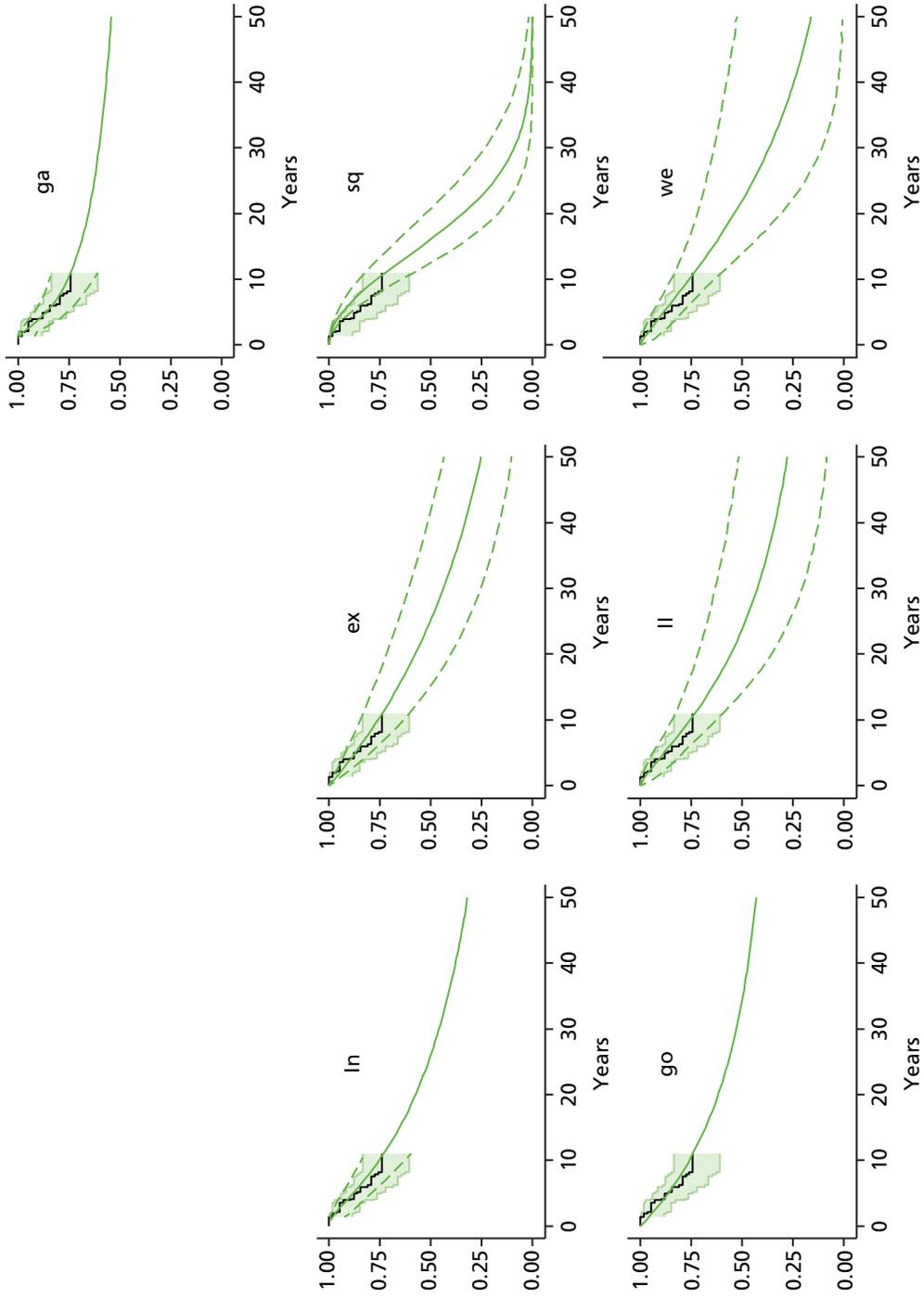


FIGURE 49 Bentley *et al.* 2012⁷⁸ ACI arm model fits. Analysis time = years. Vertical axis = proportion not failed.

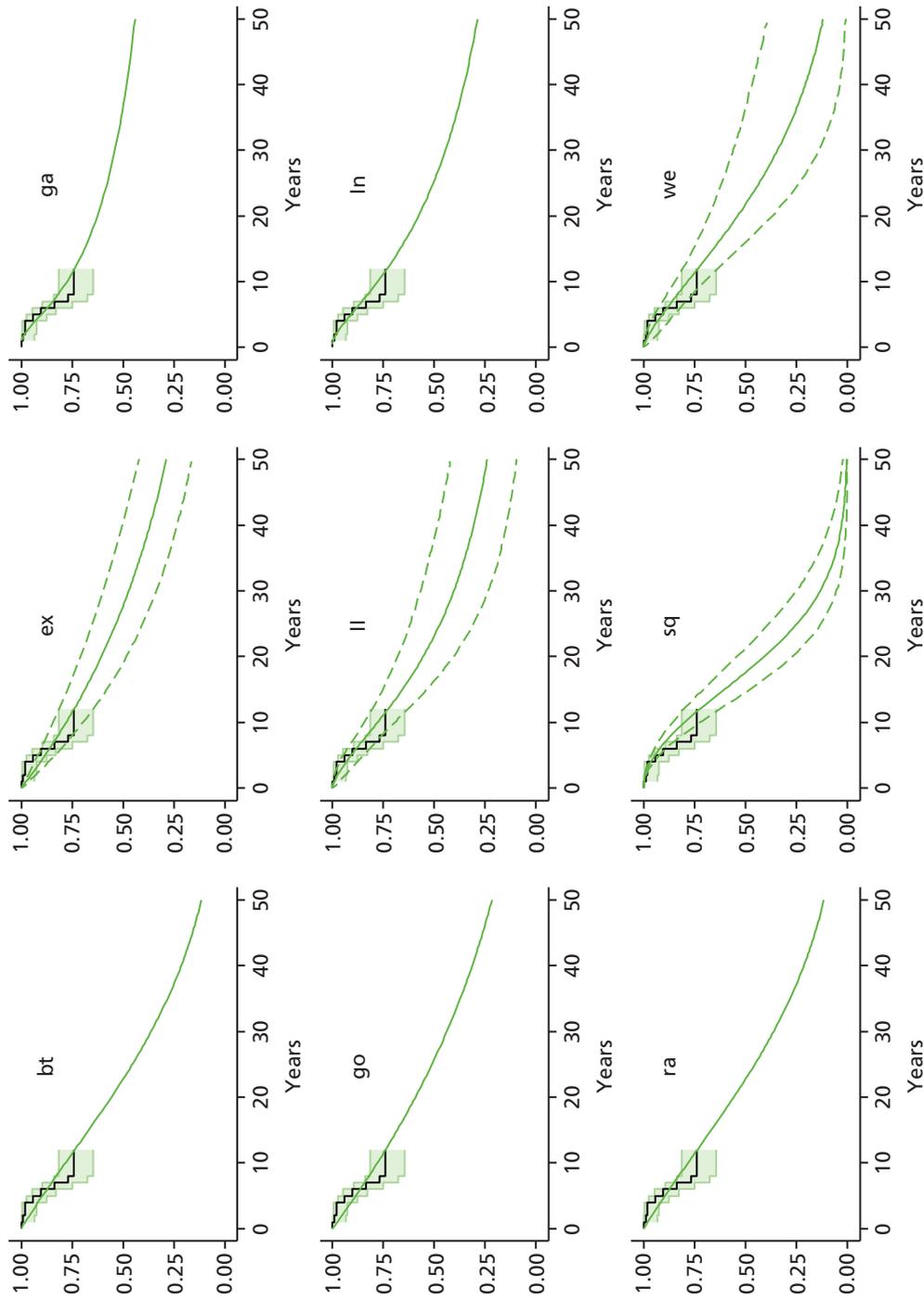


FIGURE 50 Biant *et al.* 2014⁷⁹ ACI model fits. Analysis time = years. Vertical axis = proportion not failed.

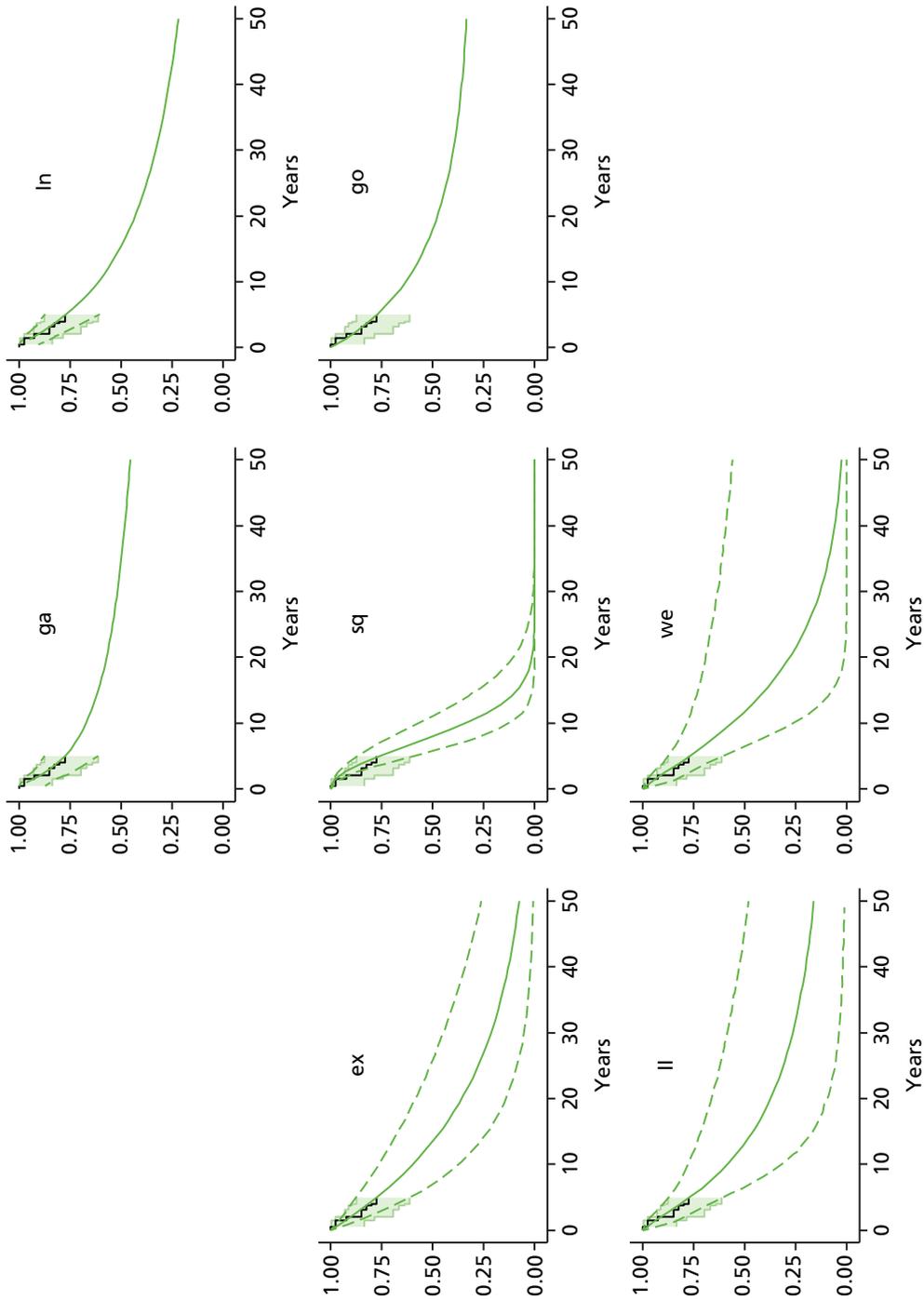


FIGURE 51 Knutsen et al. 2007⁶⁷ model fits, ACI arm. Analysis time = years. Vertical axis = proportion not failed.

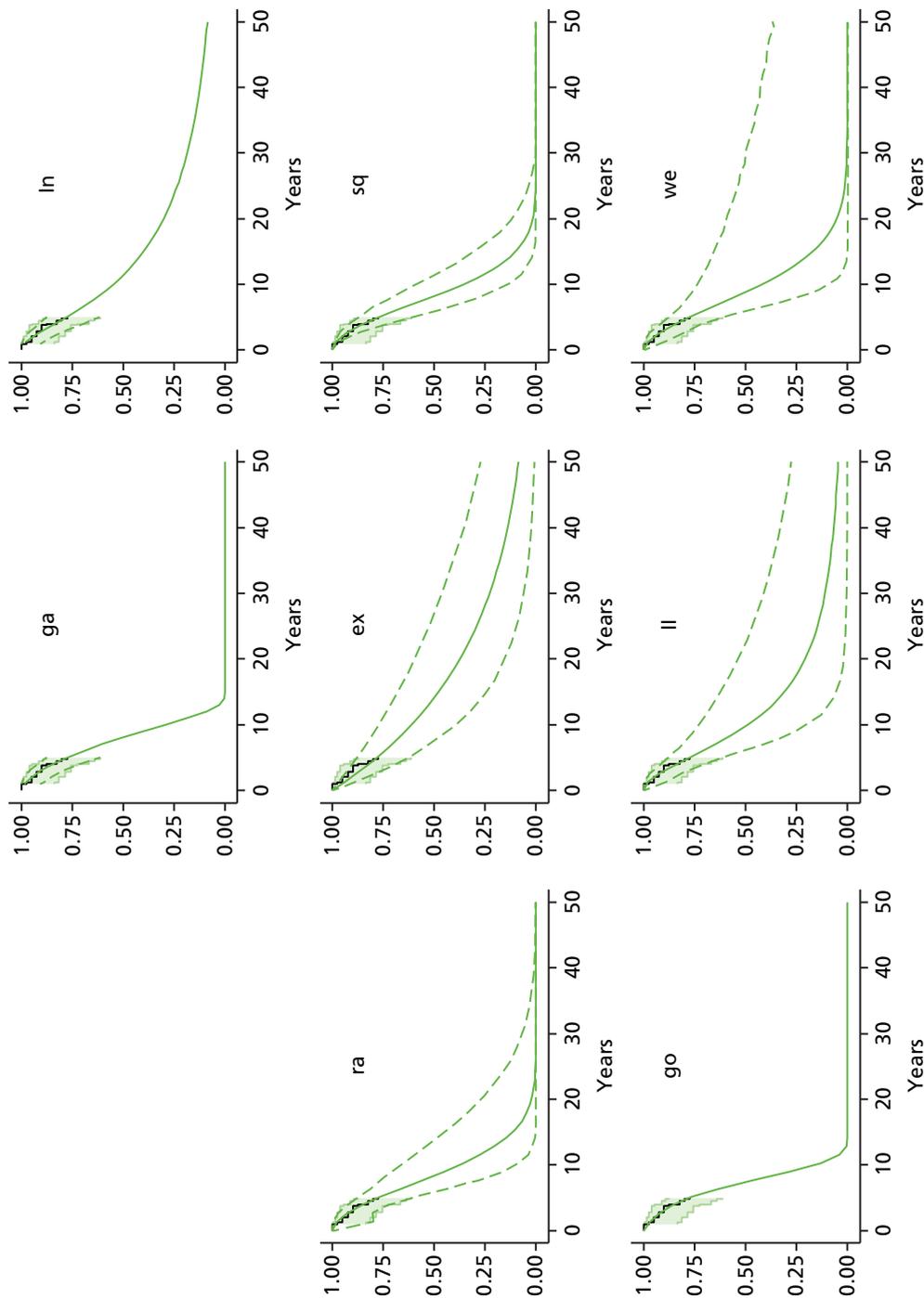


FIGURE 52 Knutsen et al. 2007⁶⁷ model fits, MIF arm. Analysis time = years. Vertical axis = proportion not failed.

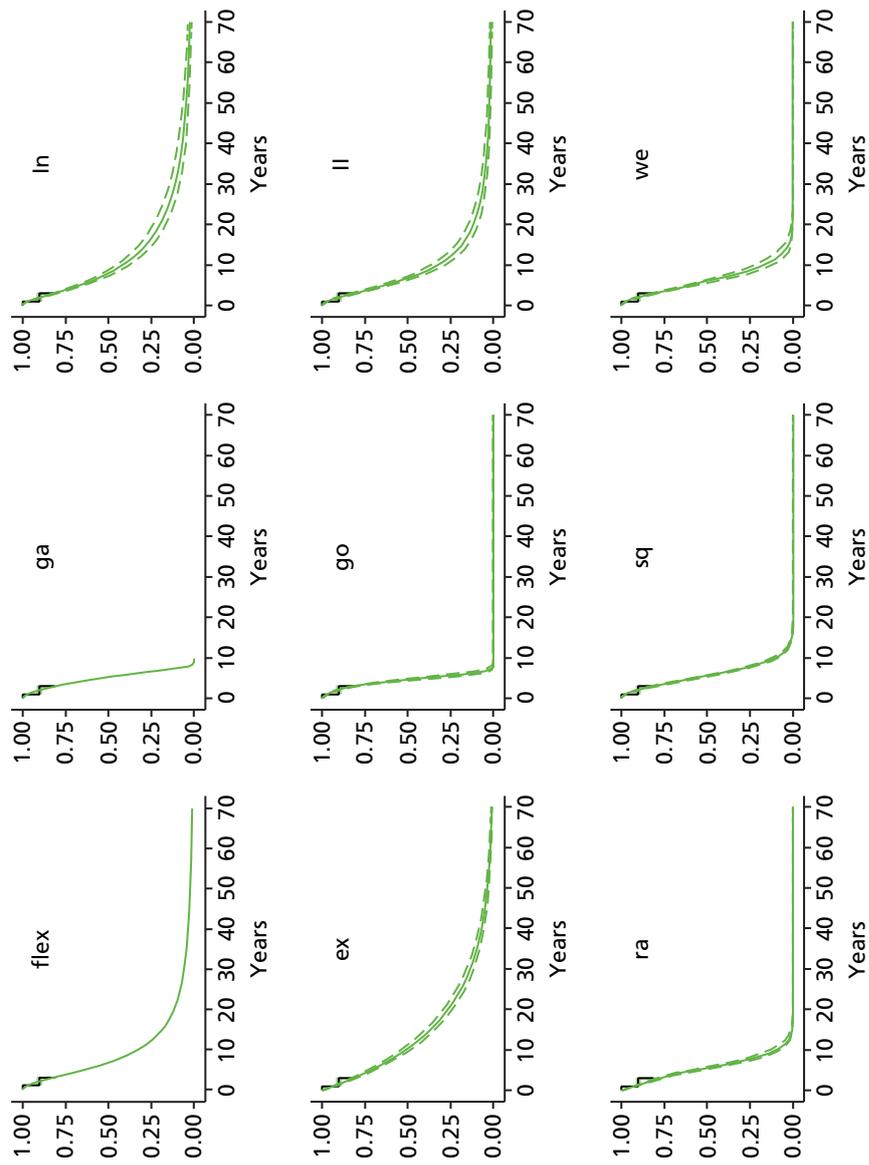


FIGURE 53 Layton et al. 2015¹¹³ model fits, MF. Vertical axis = proportion not failed.

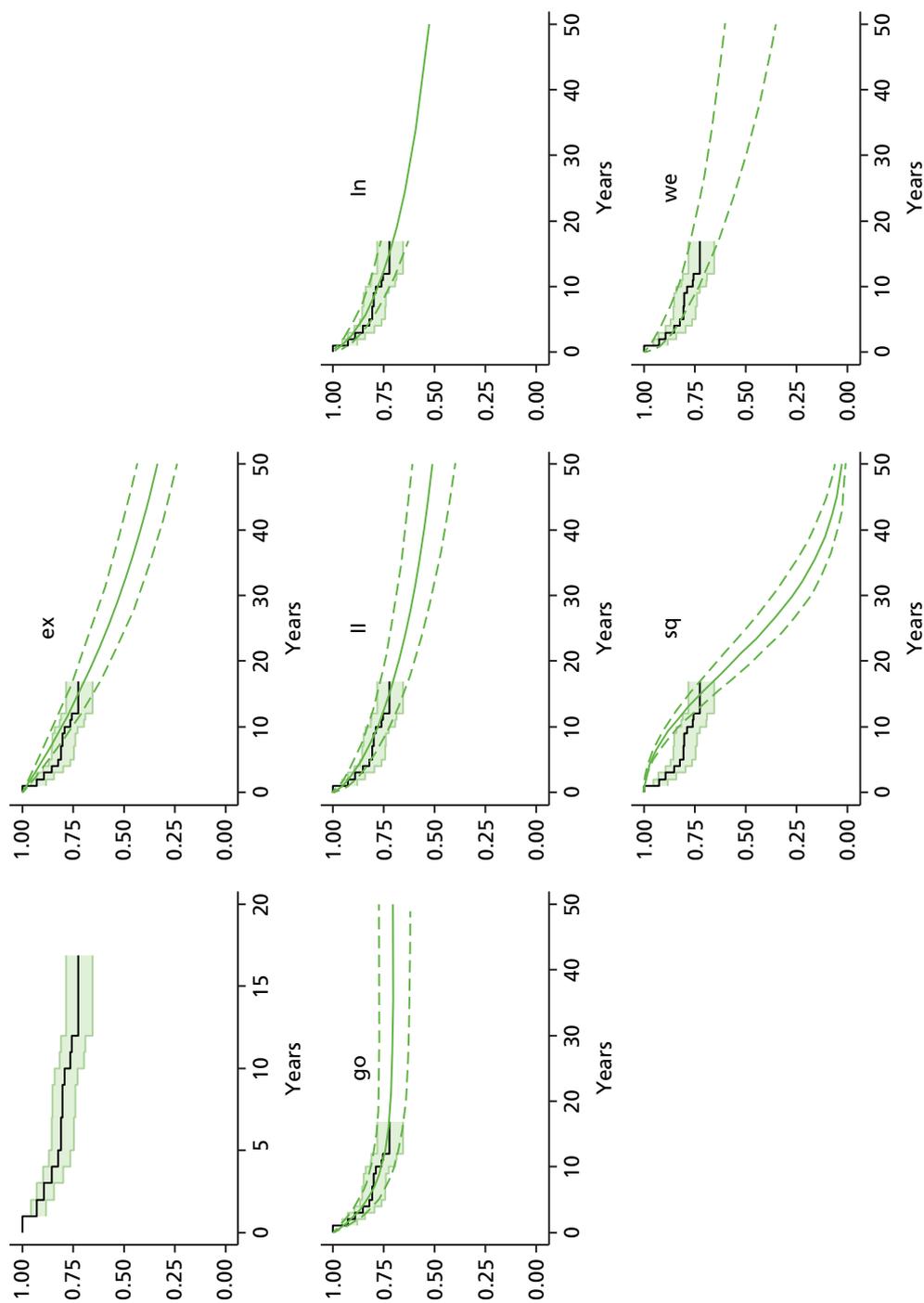


FIGURE 54 Minas *et al.* 2014¹⁶⁴ model fits, ACI. Analysis time = years. Vertical axis = proportion not failed.

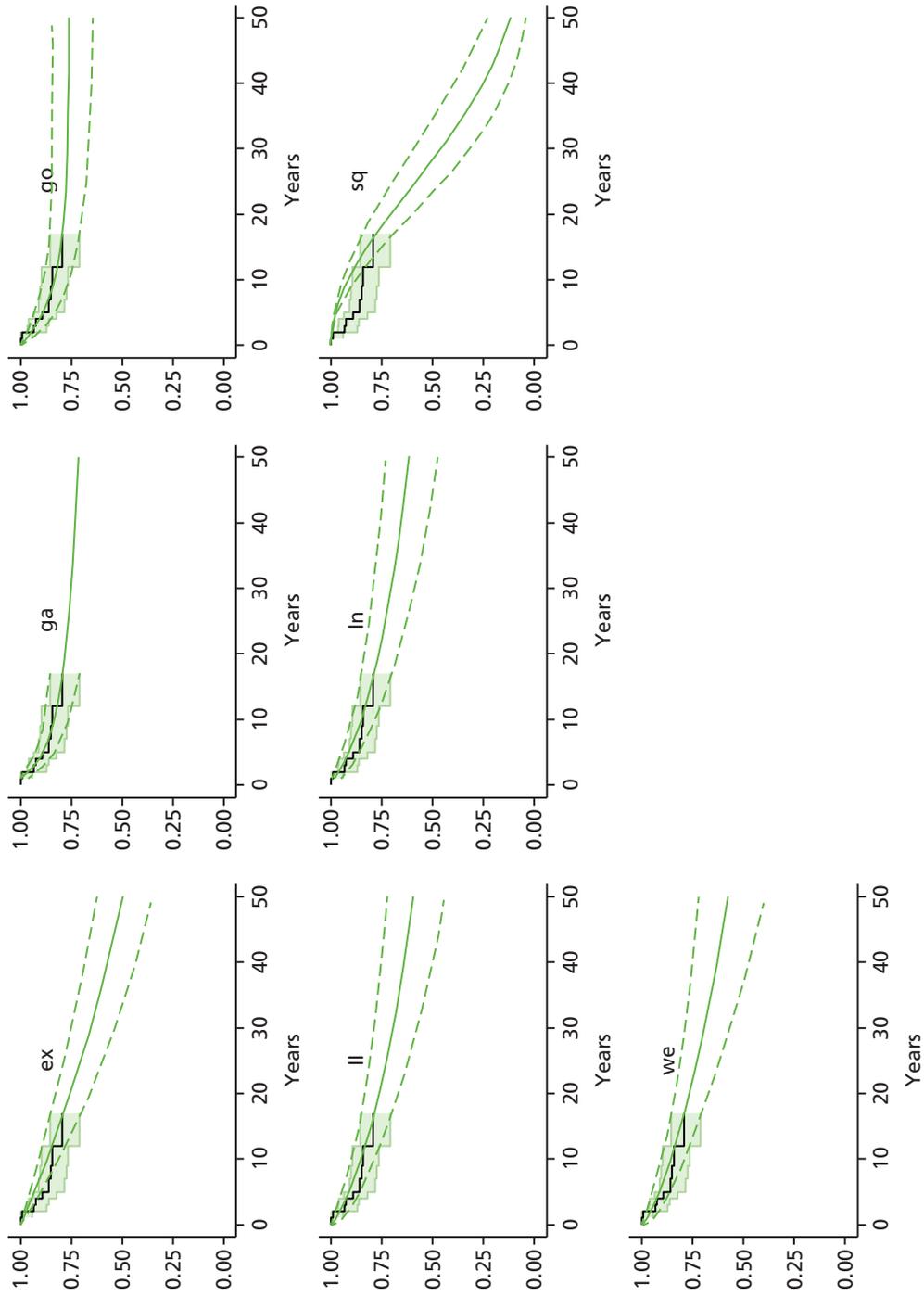


FIGURE 55 Minas et al. 2014⁶⁴ model fits, ACI, no previous intervention. Analysis time = years. Vertical axis = proportion not failed.

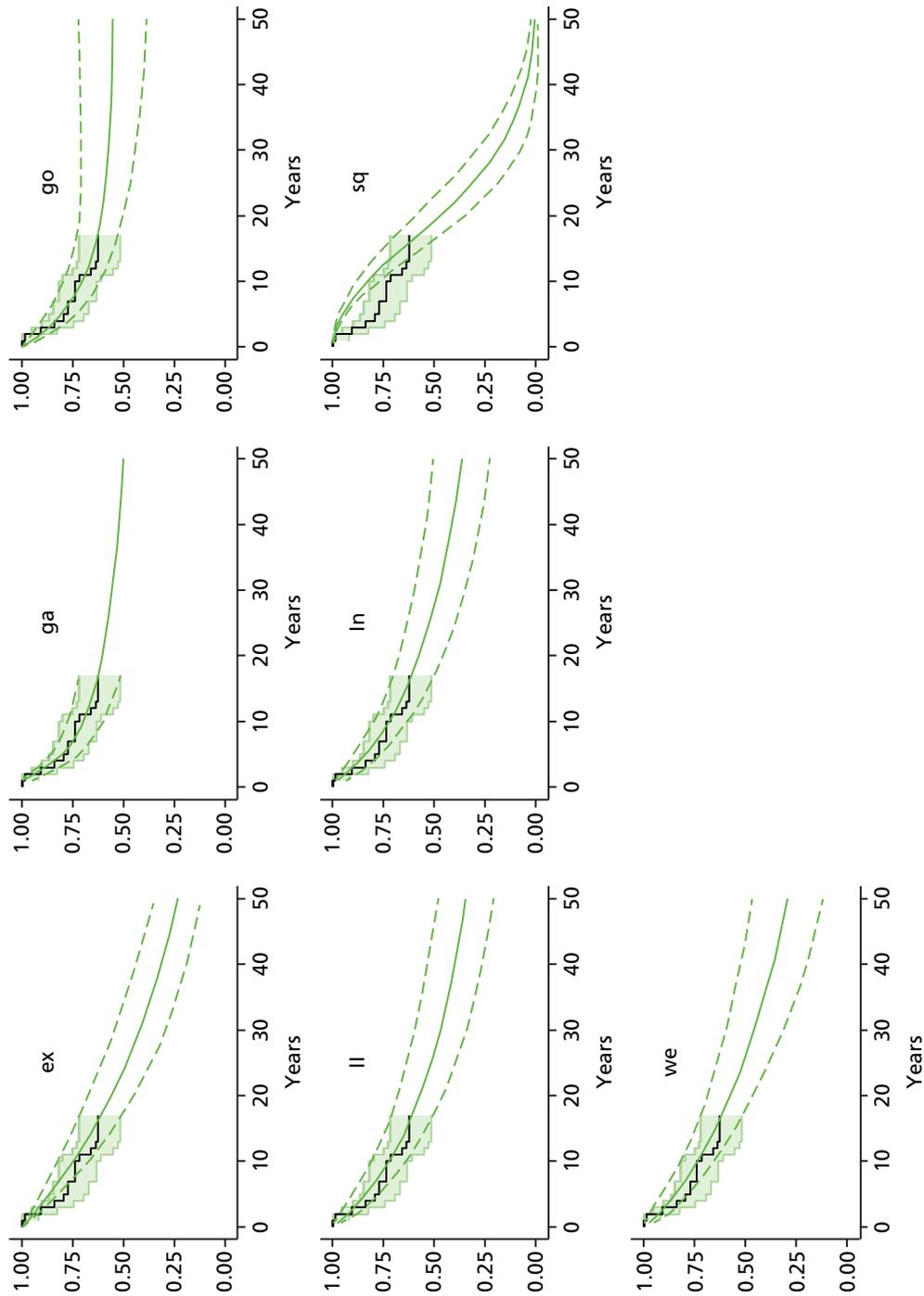


FIGURE 56 Minas *et al.* 2014¹⁶⁴ model fits, ACI, previous intervention. Analysis time = years. Vertical axis = proportion not failed.

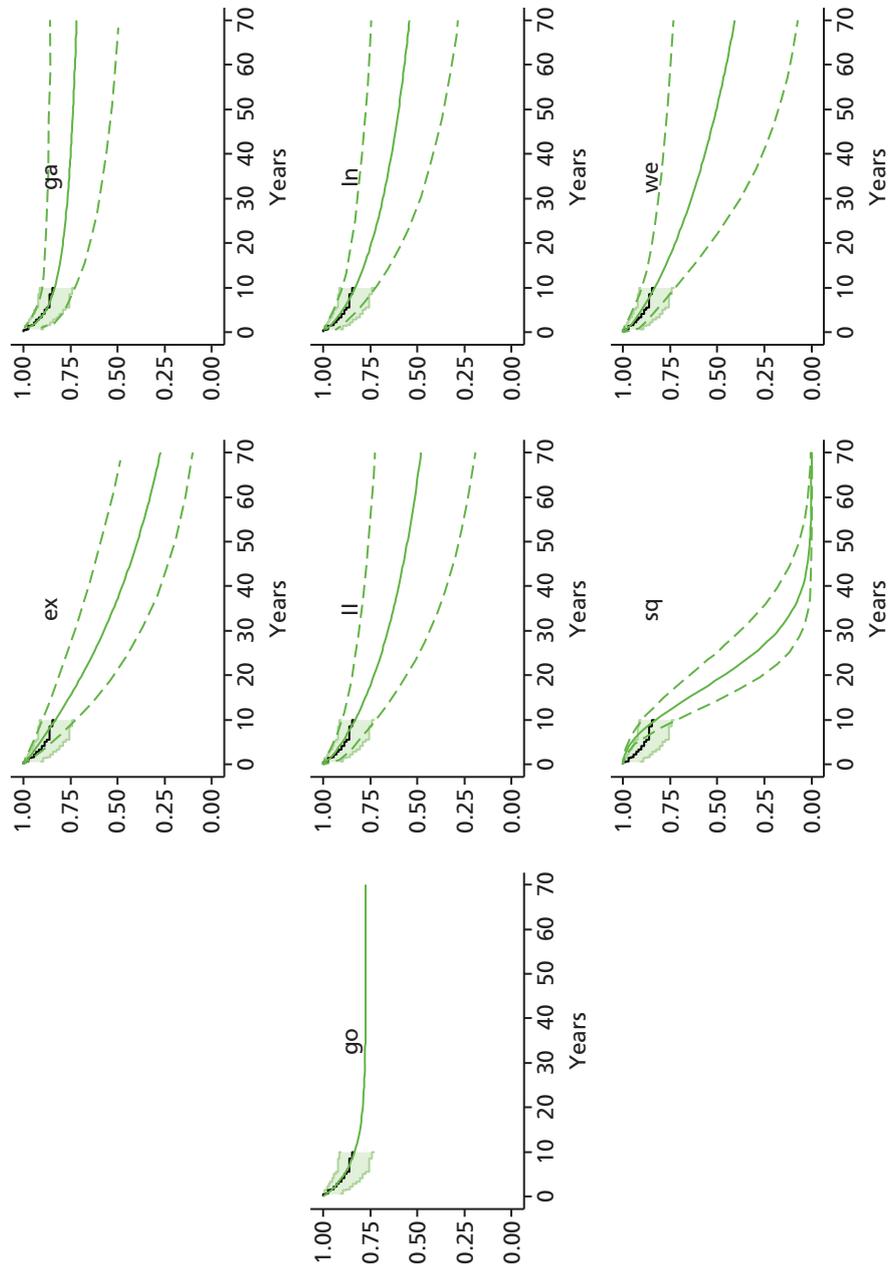


FIGURE 57 Moseley et al. 2010¹¹⁵ model fits, ACI. Analysis time = years. Vertical axis = proportion not failed.

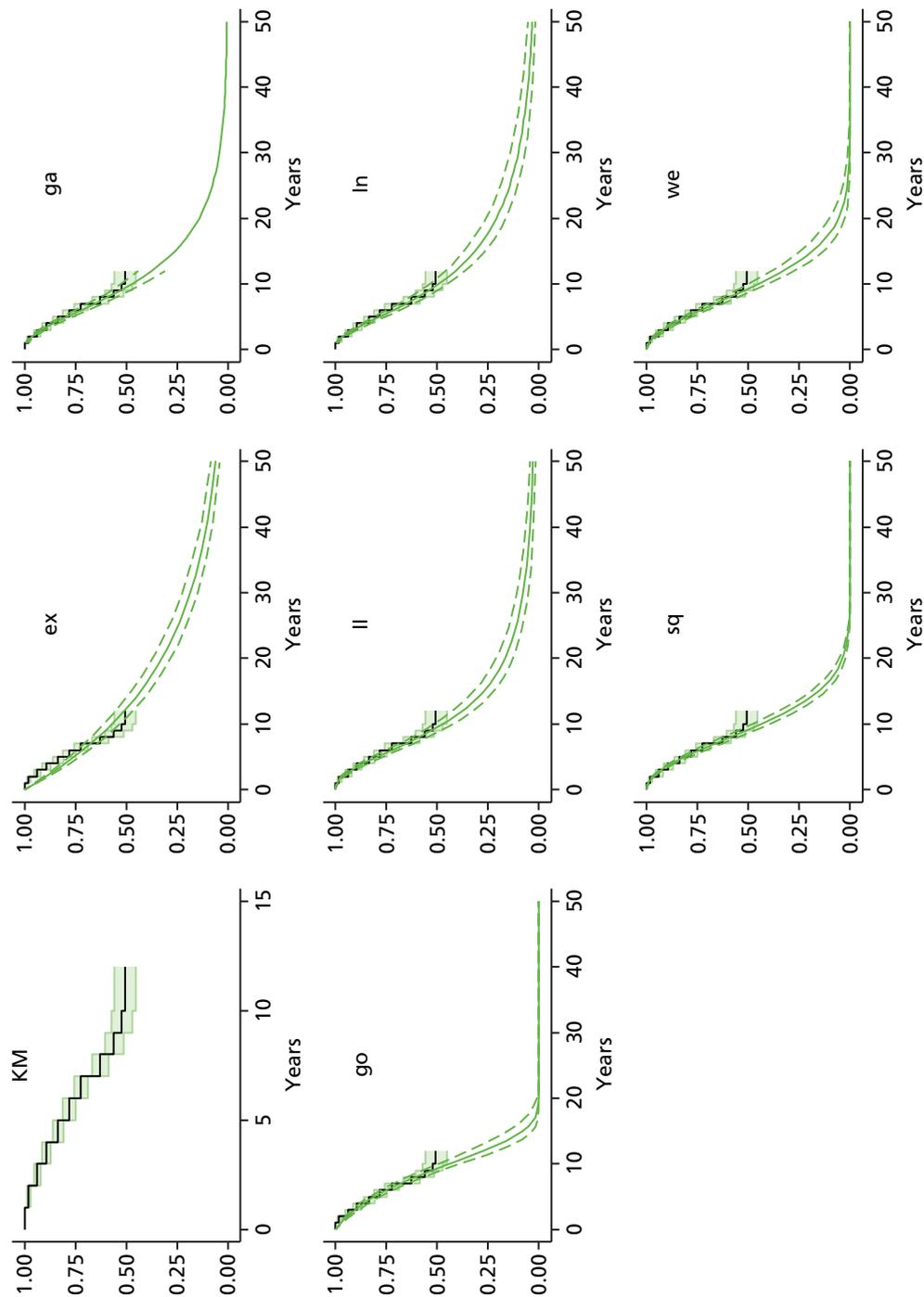


FIGURE 58 Nawaz *et al.* 2014⁶⁰ model fits, ACI, whole cohort. Analysis time = years. Vertical axis = proportion not failed.

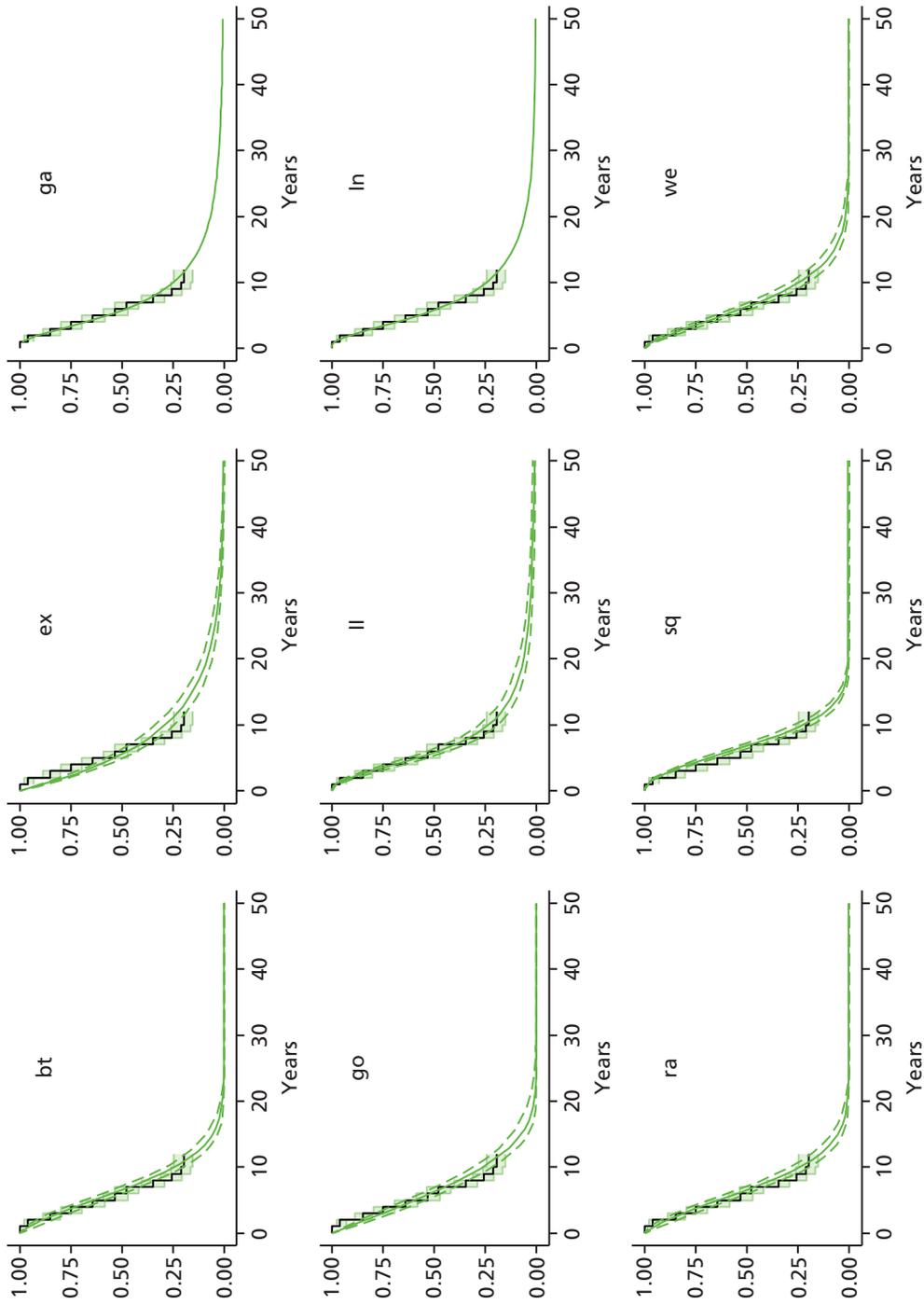


FIGURE 59 Nawaz et al. 2014⁸⁰ model fits, ACI, previous intervention. Analysis time = years. Vertical axis = proportion not failed.

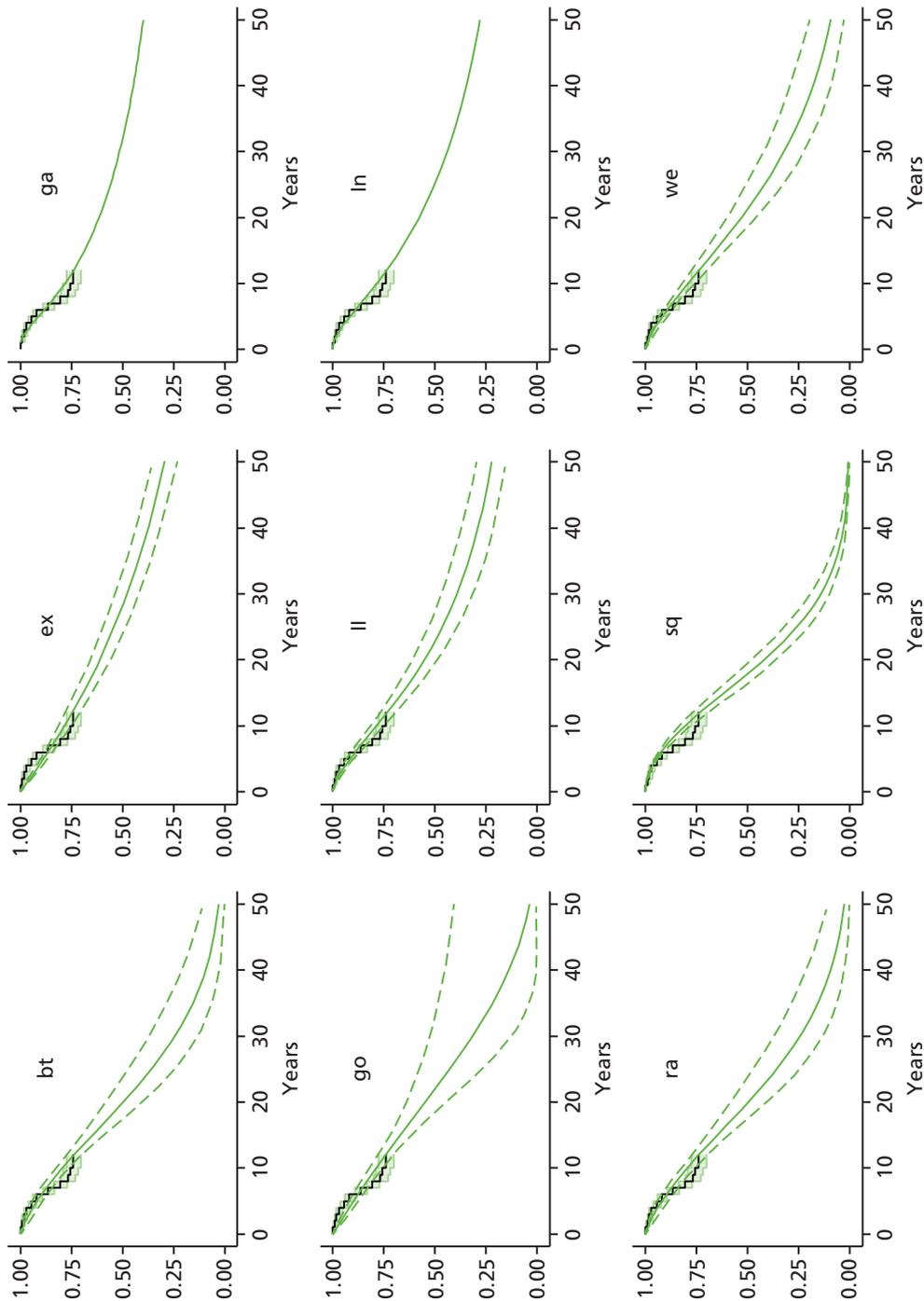


FIGURE 60 Nawaz *et al.* 2014⁸⁰ model fits, ACI, no previous intervention. Analysis time = years. Vertical axis = proportion not failed.

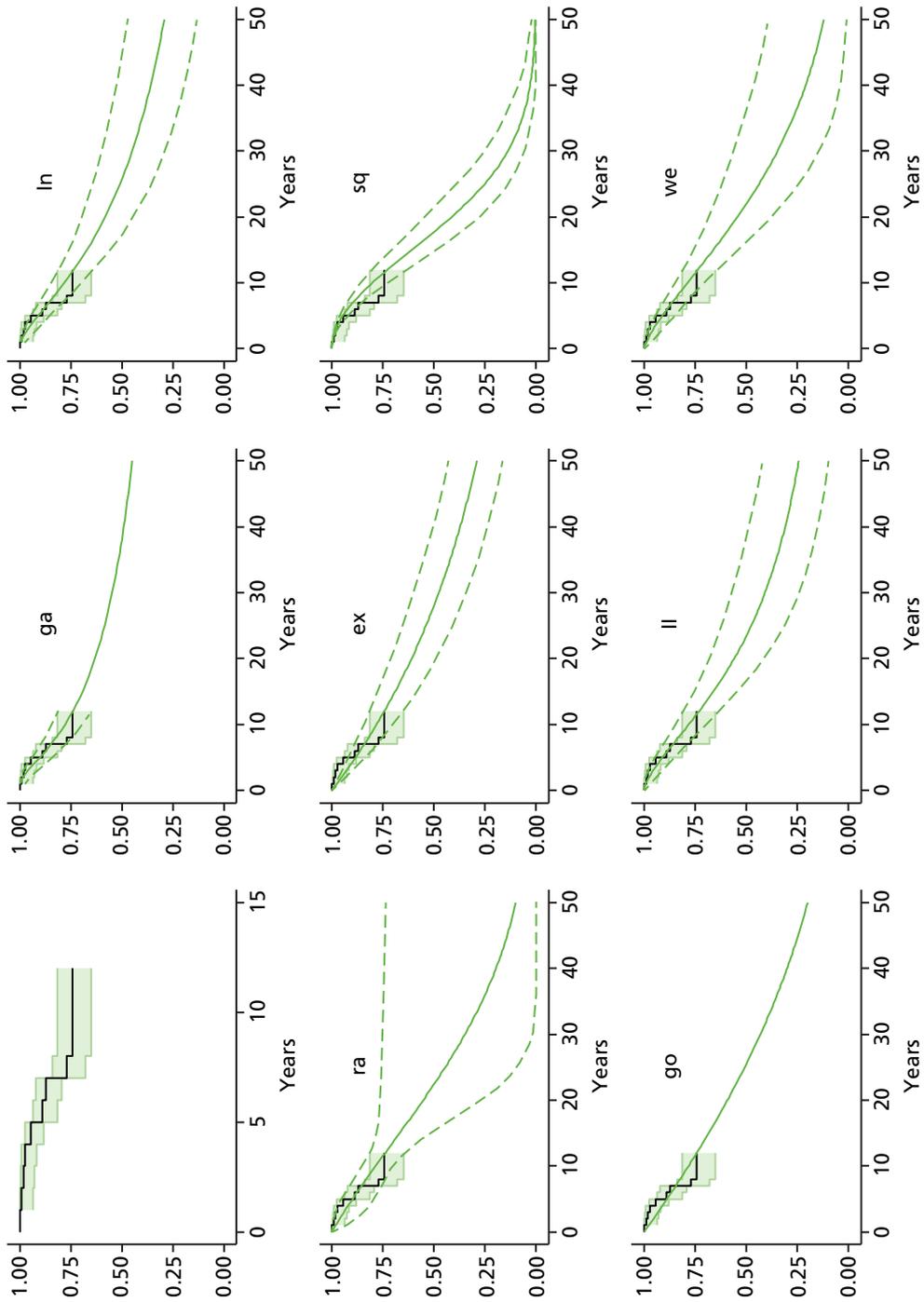


FIGURE 61 Nawaz et al. 2014⁸⁰ model fits, ACI, lateral femoral site. Analysis time = years. Vertical axis = proportion not failed.

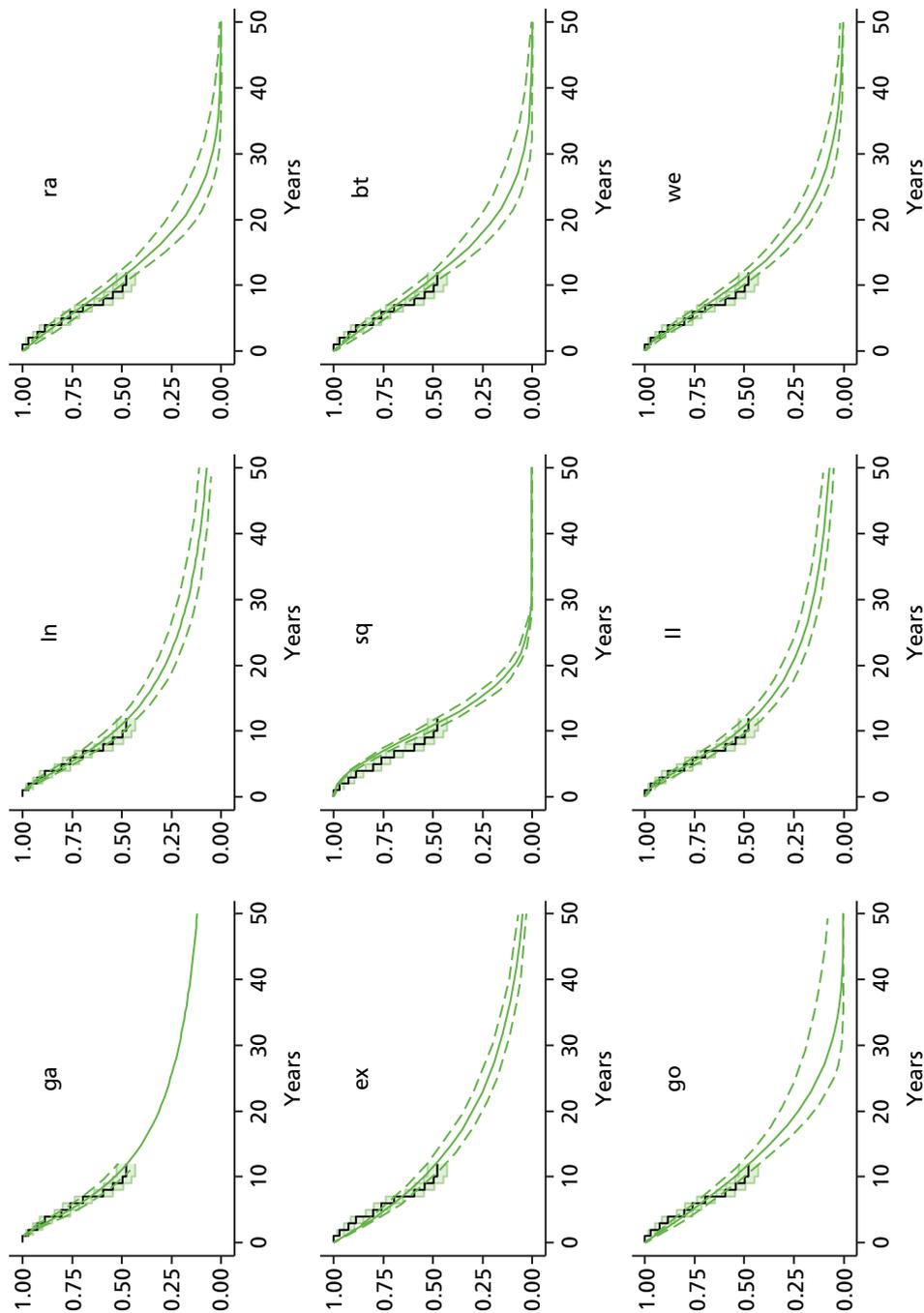


FIGURE 62 Nawaz *et al.* 2014⁸⁰ model fits, ACI, medial femoral site. Analysis time = years. Vertical axis = proportion not failed.

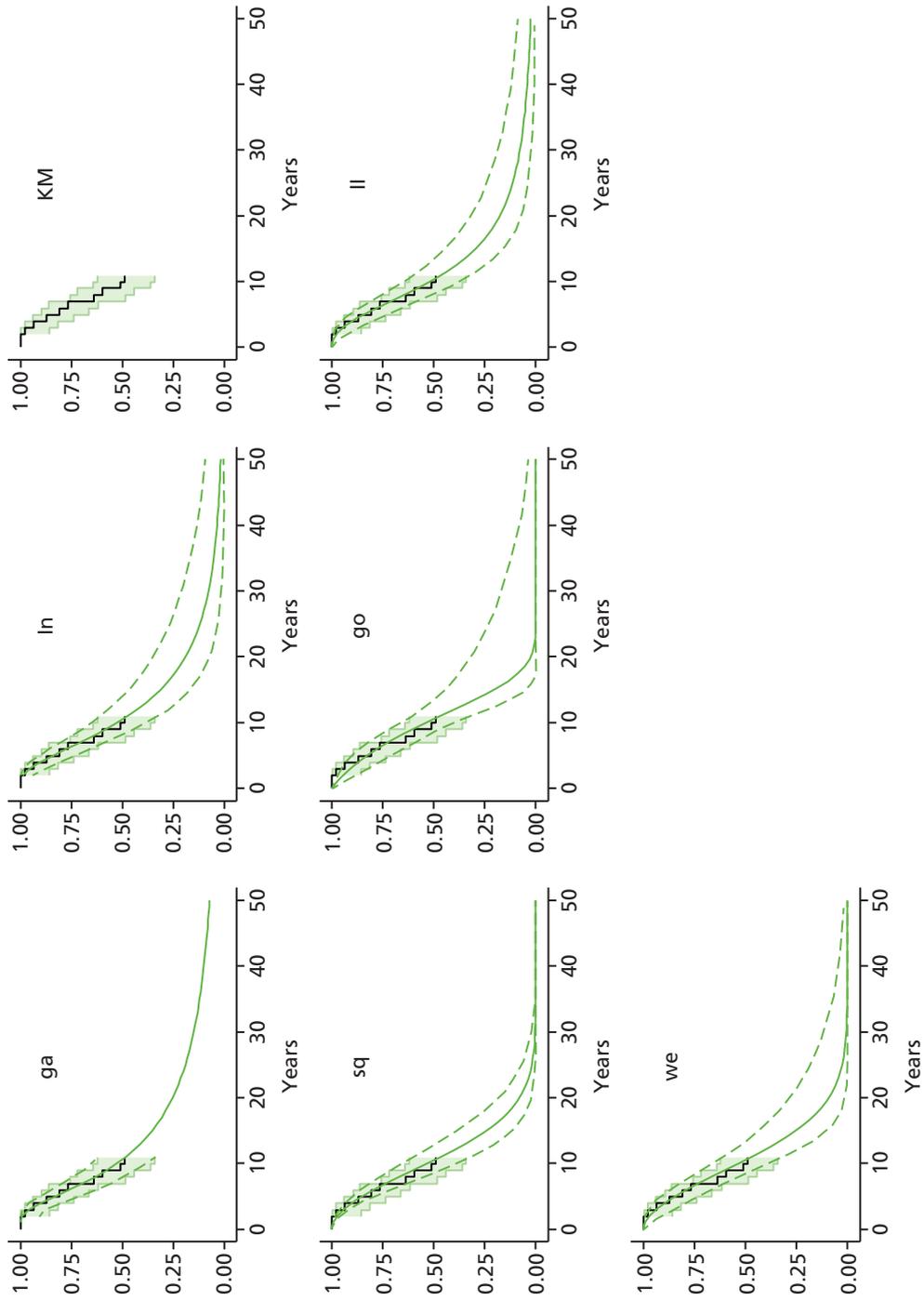


FIGURE 63 Nawaz et al. 2014⁸⁰ model fits, ACI, multisite. Analysis time = years. Vertical axis = proportion not failed.

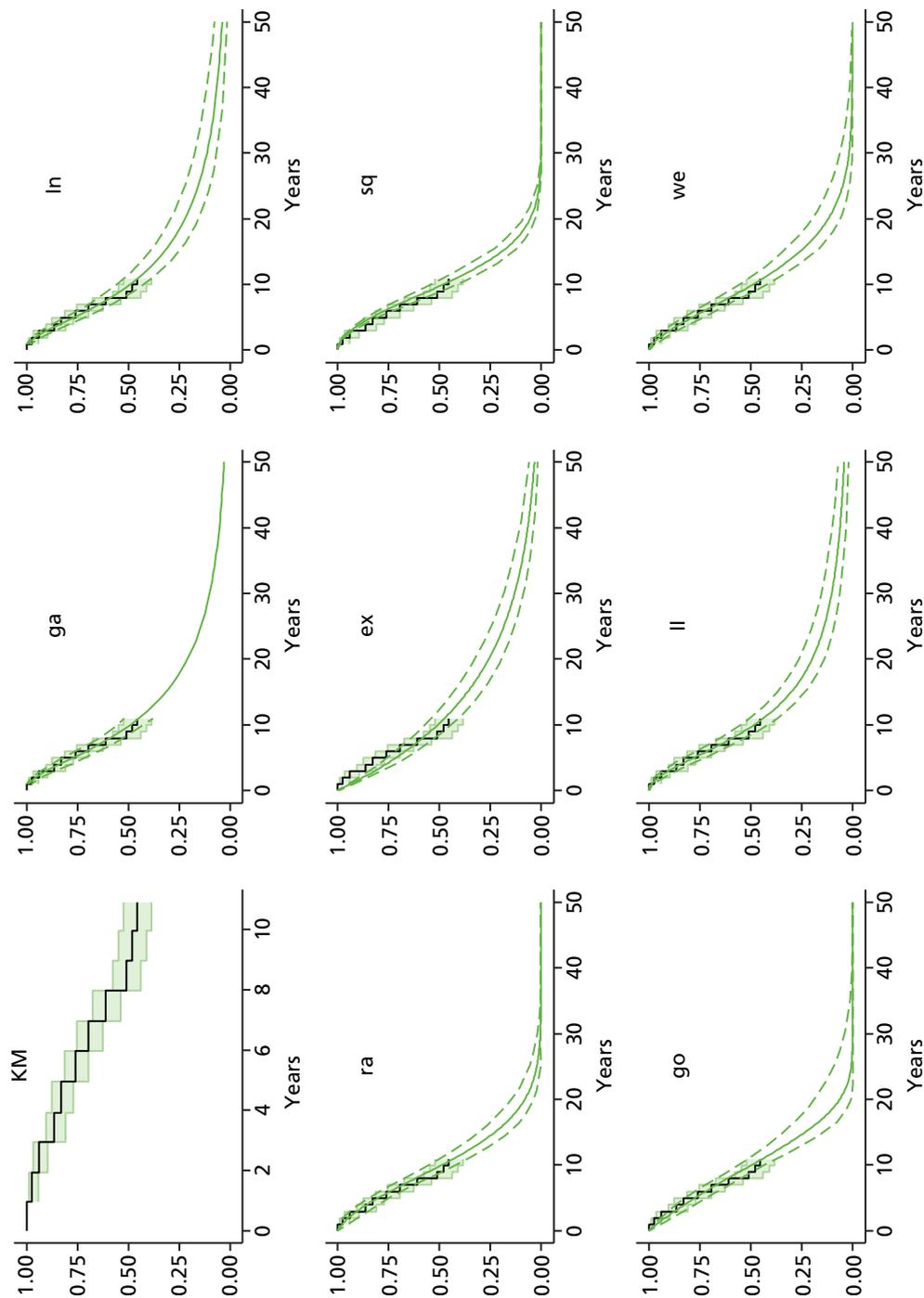


FIGURE 64 Nawaz *et al.* 2014⁸⁰ model fits, ACI, patella site. Analysis time = years. Vertical axis = proportion not failed.

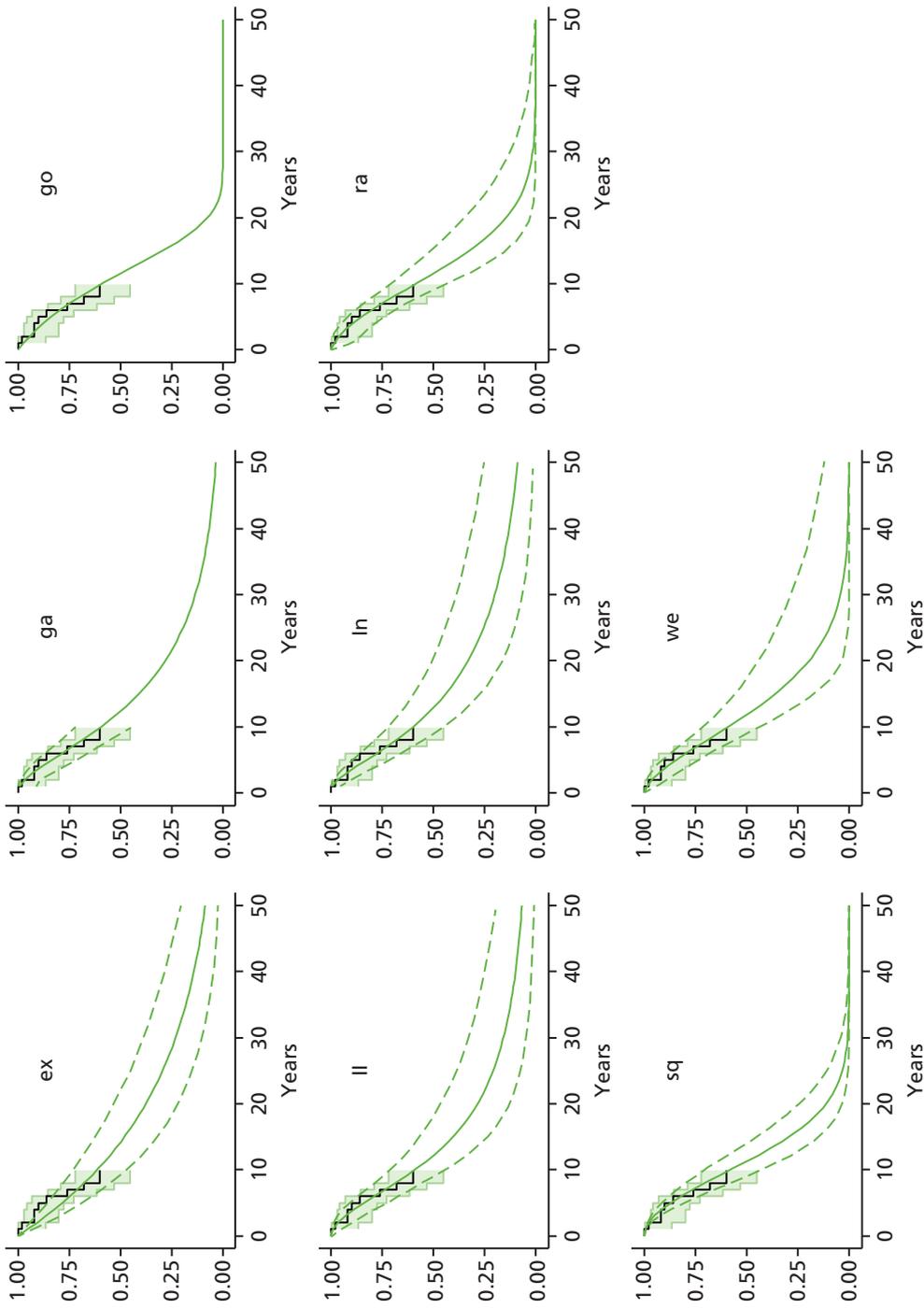


FIGURE 65 Nawaz et al. 2014⁸⁰ model fits, ACI, trochlea site. Analysis time = years. Vertical axis = proportion not failed.

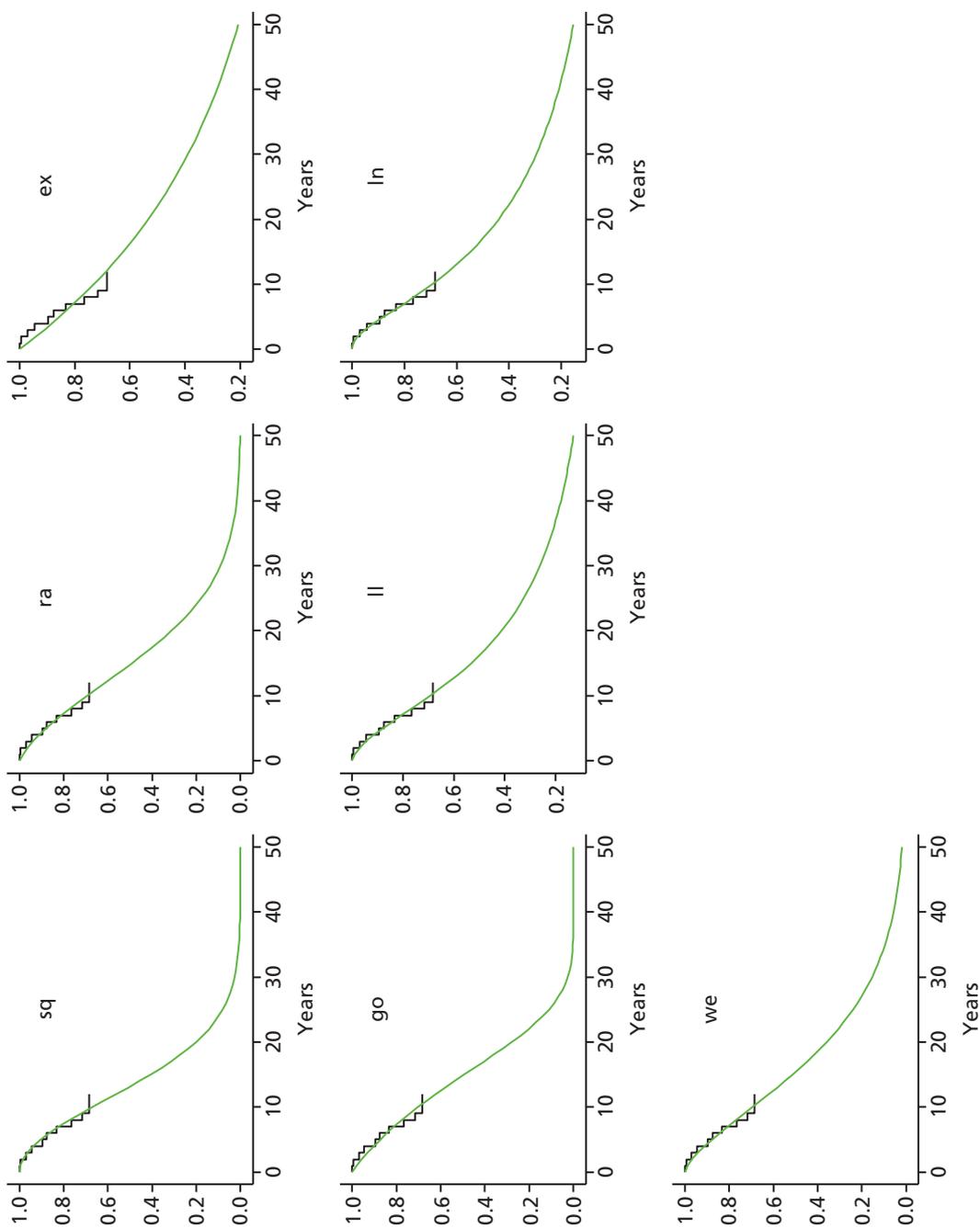


FIGURE 66 Nawaz *et al.* 2014⁸⁰ model fits, ACI, grade 0 degradative change. Analysis time = years. Vertical axis = proportion not failed.

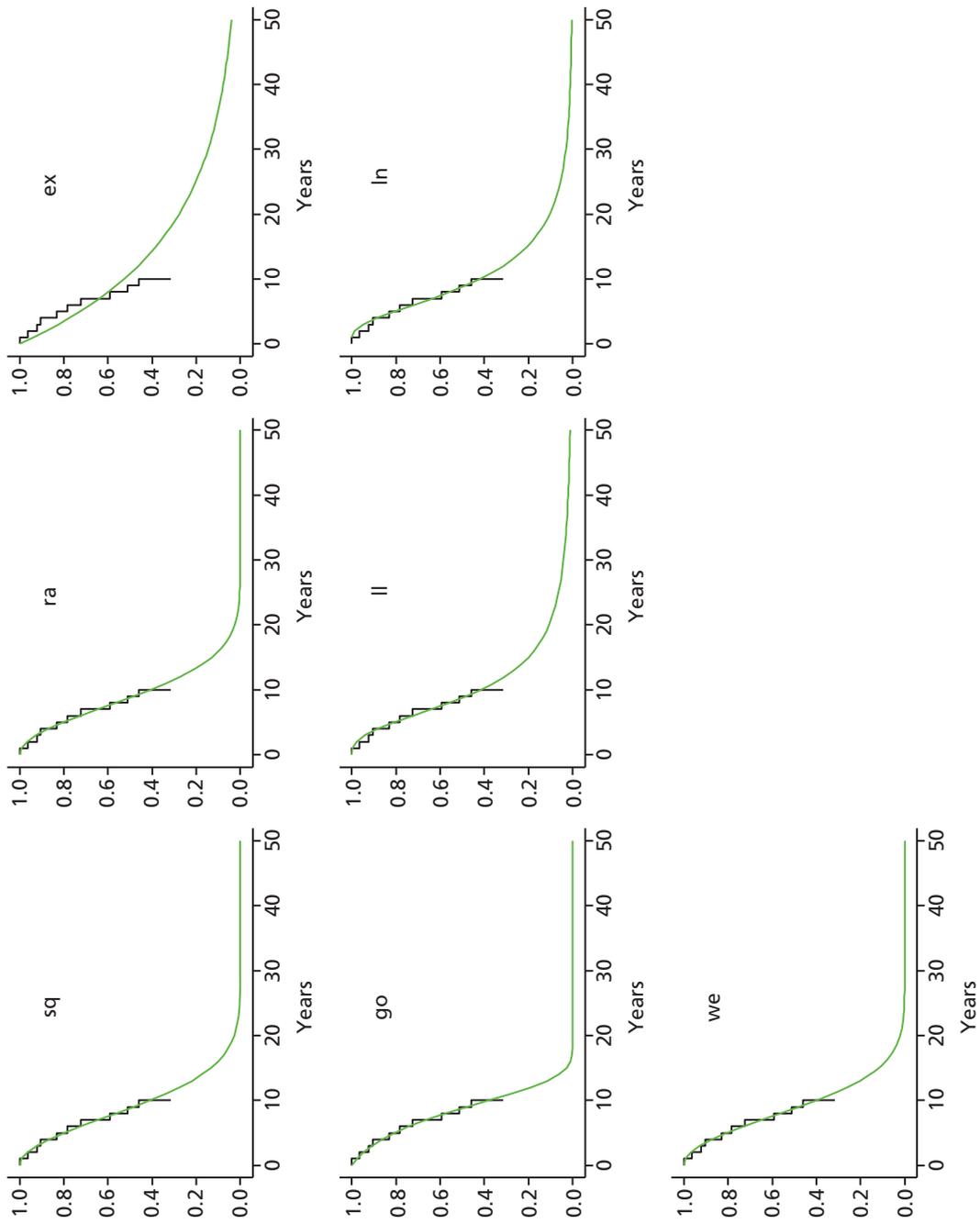


FIGURE 67 Nawaz et al. 2014⁸⁰ model fits, ACI, grade 1 degradative change. Analysis time = years. Vertical axis = proportion not failed.

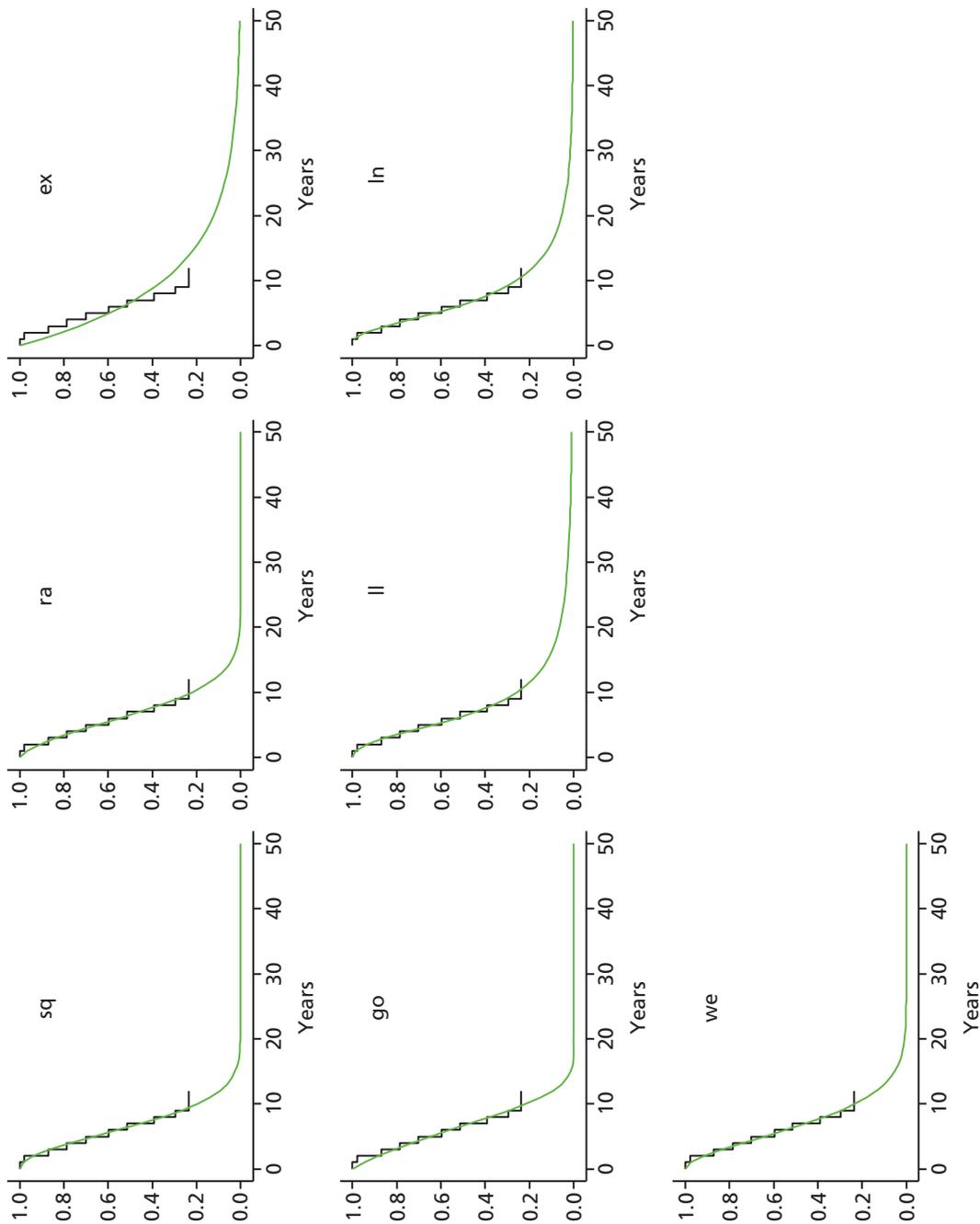


FIGURE 68 Nawaz *et al.* 2014⁸⁰ model fits, ACI, grade 2 degradative change. Analysis time = years. Vertical axis = proportion not failed.

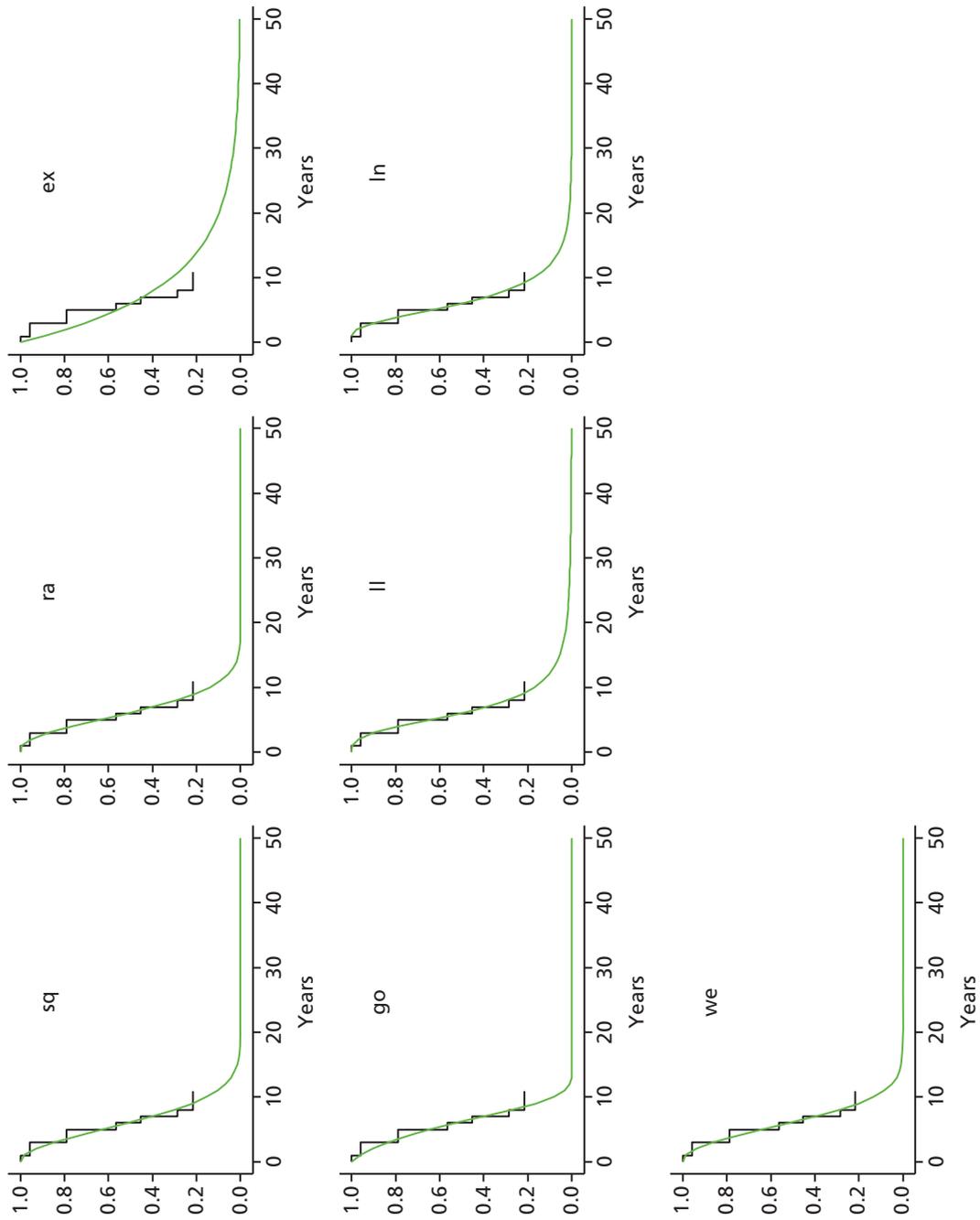


FIGURE 69 Nawaz et al. 2014⁸⁰ model fits, ACI, grade 3 degradative change. Analysis time = years. Vertical axis = proportion not failed.

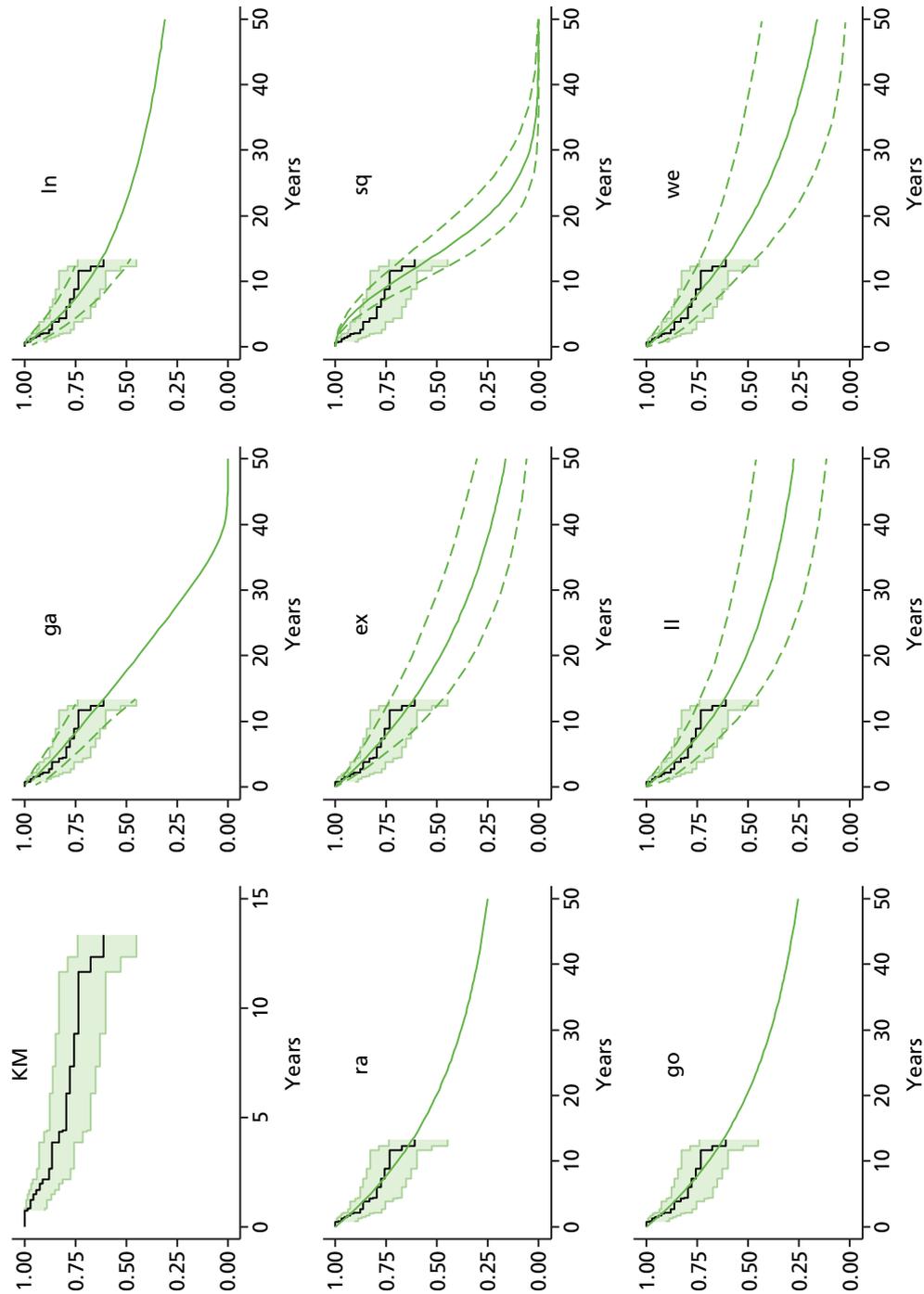


FIGURE 70 Niemeyer et al. 2014¹¹⁷ model fits, ACI.

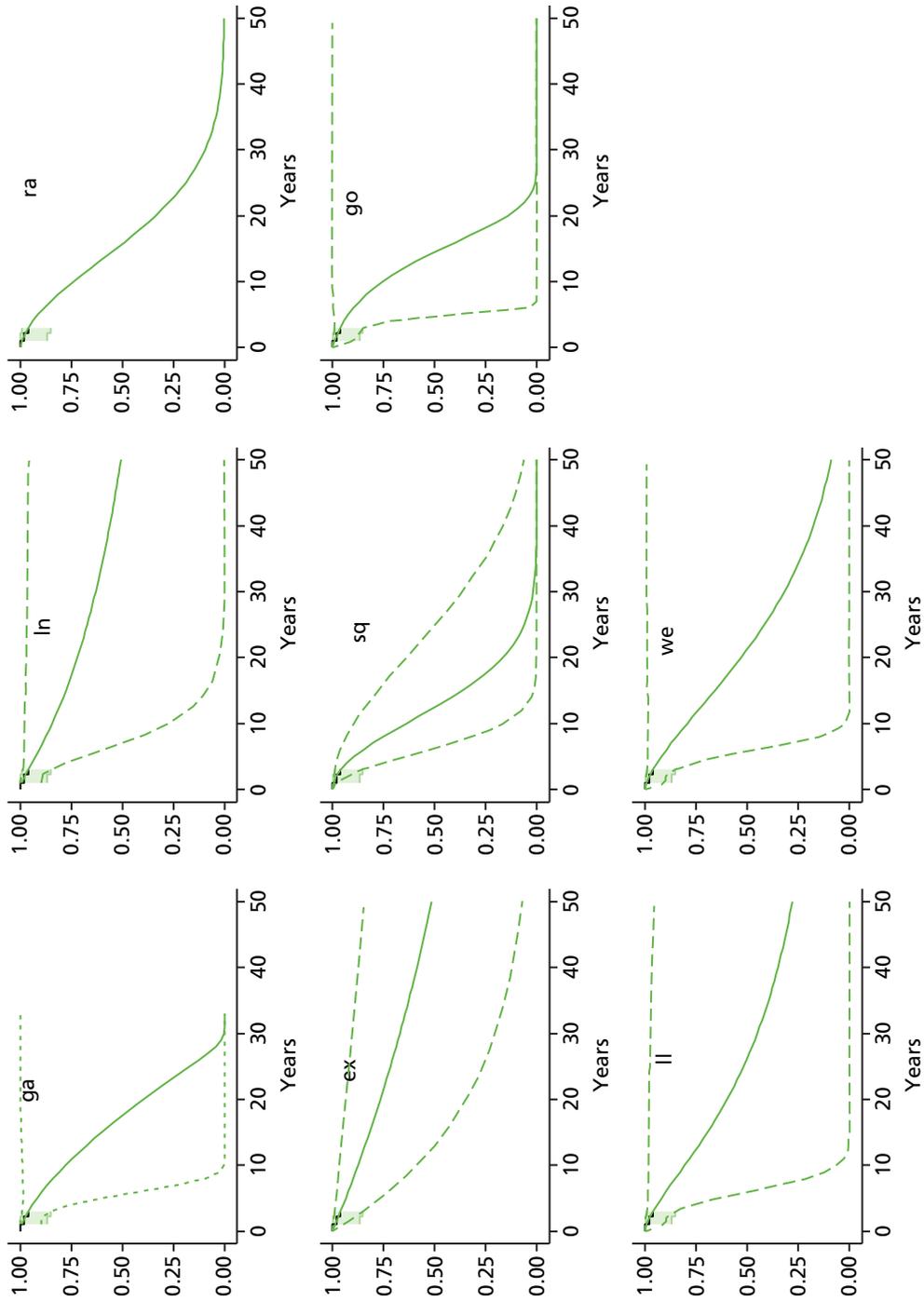


FIGURE 71 Saris et al. 2009,⁷⁰ ACI arm.

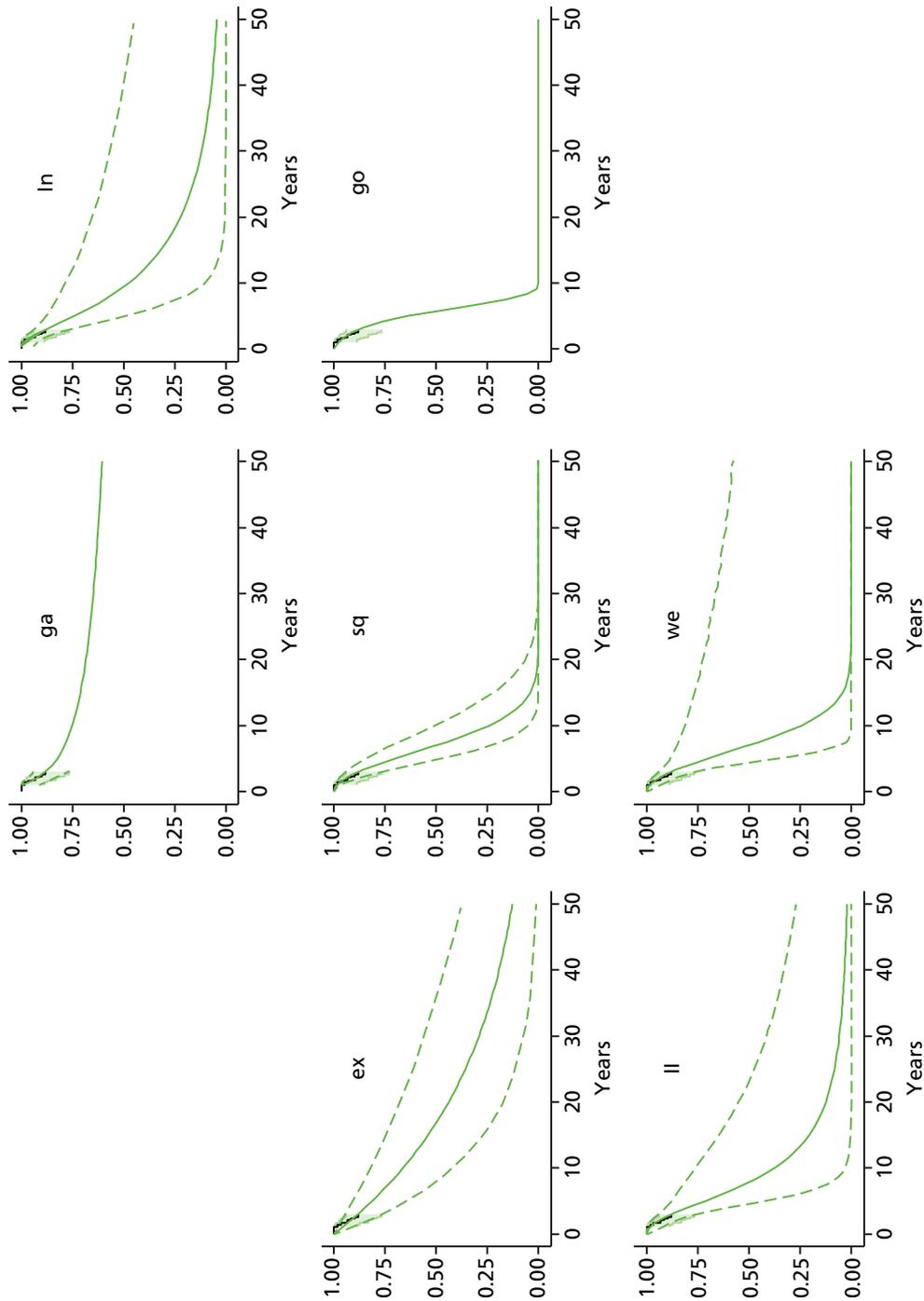


FIGURE 72 Saris *et al.* 2009,⁷⁰ MF arm.

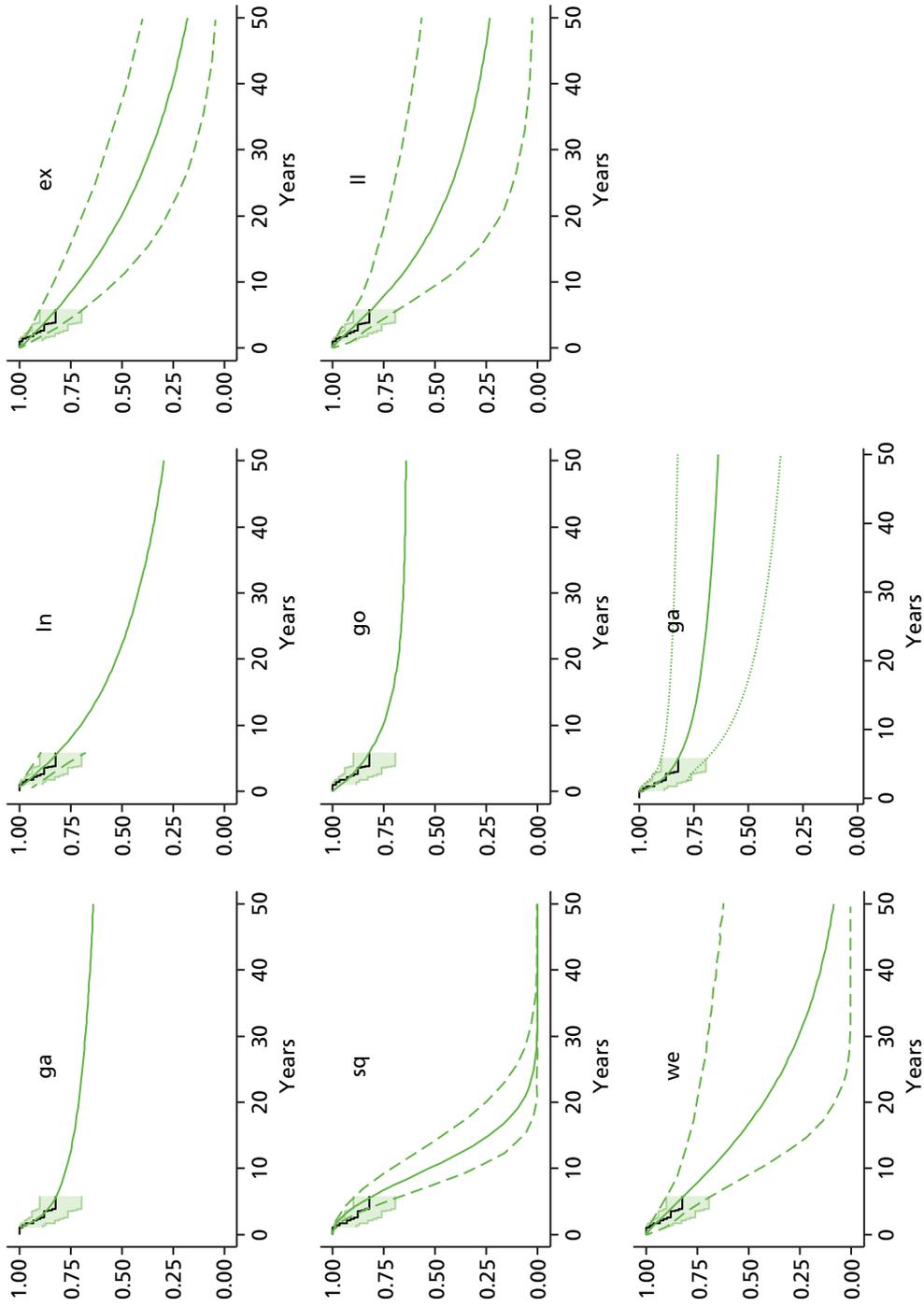


FIGURE 73 Vanlauwe et al. 2011,⁴³ MF arm model fits; IPD reconstructed without risk table data.

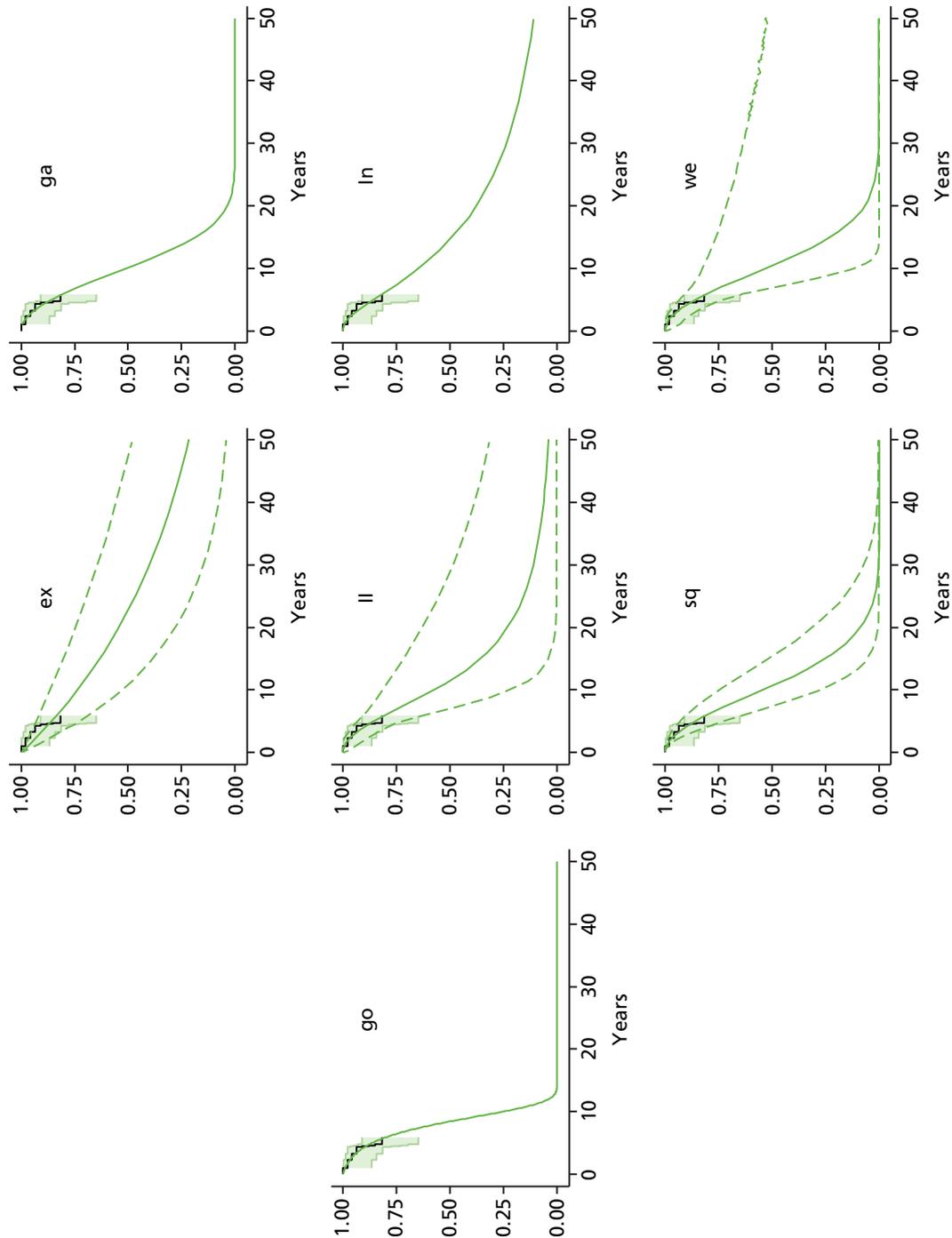


FIGURE 74 Vanlauwe *et al.* 2011,⁴³ ACI arm model fits.

Appendix 10 Economic search strategies

MEDLINE search strategy (1946 to July 2014)

1. exp Economics/
2. exp "Costs and Cost Analysis"/
3. exp Cost-Benefit Analysis/
4. Health Status/
5. exp "Quality of Life"/
6. exp Quality-Adjusted Life Years/
7. (pharmacoeconomic* or pharmaco-economic* or economic* or cost*).tw.
8. (health state* or health status).tw.
9. (qaly* or ICER* or utilit* or EQ5D or EQ-5D or euroqol or euro-qol or short-form 36 or short form 36 or SF-36 or SF36 or SF-6D or SF6D or SF-12 or SF12 or health utilities index or HUI).tw.
10. (markov or time trade off or TTO or standard gamble or SG or hrql or hrqol or disabilit* or disutilit* or net benefit or net-benefit or contingent valuation).tw.
11. (quality adj2 life).tw.
12. (decision adj2 model).tw.
13. (quality of wellbeing or qwb visual analog* scale* or discrete choice experiment* or health* year* equivalen* or hyes or hye or 15-D or 15D or (willing* adj2 pay)).tw.
14. ("resource use" or resource utili?ation or resource\$).tw.
15. (utility* adj2 (value* or index* or health or measure* or estimate*)).tw.
16. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15
17. exp Chondrocytes/tr [Transplantation]
18. exp Cartilage, Articular/tr [Transplantation]
19. exp Transplantation, Autologous/
20. (MACI or MACT or chondroelect or ACI).tw.
21. (chondrocyte* adj4 (implant* or transplant* or matrix or characteri*)).tw.
22. (autologous adj3 (implant* or transplant* or chondrocyte* or cartilage)).tw.
23. (cartilage* adj2 (transplant* or implant*)).tw.
24. 17 or 18 or 19 or 20 or 21 or 22 or 23
25. Knee/ or knee*.mp.
26. 24 and 25
27. Animals/
28. Humans/
29. 27 not 28
30. 26 not 29
31. 16 and 30
32. limit 31 to yr="2004 -Current"

EMBASE search strategy (1947 to July 2014)

1. exp health economics/
2. exp health status/
3. exp "quality of life"/
4. exp quality adjusted life year/
5. (pharmacoeconomic* or pharmaco-economic* or economic* or cost*).tw.
6. (health state* or health status).tw.
7. (qaly* or ICER* or utilit* or EQ5D or EQ-5D or euroqol or euro-qol or short-form 36 or SF-36 or SF36 or SF-12 or SF12 or SF36 or SF-6D or SF6D or health utilities index or HUI).tw.

8. (markov or time trade off or TTO or standard gamble or SG or hrql or hrqol or disabilit* or disutilit* or net benefit* or contingent valuation).tw.
9. (quality adj2 life).tw.
10. (decision adj2 model).tw.
11. ("quality of wellbeing" or "quality of well-being" or qwb or visual analog* scale* or discrete choice experiment* or health* year* equivalen* or hye* or (willing* adj2 pay)).tw.
12. resource*.tw.
13. (utility* adj2 (value* or index* or health or measure* or estimate*)).tw.
14. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13
15. exp *chondrocyte implantation/
16. (MACI or MACT or chondrocelect or ACI).tw.
17. (chondrocyte* adj4 (implant* or transplant* or matrix or characteri*)).tw.
18. (autologous adj3 (implant* or transplant* or chondrocyte* or cartilage)).tw.
19. (cartilage* adj2 (transplant* or implant*)).tw.
20. 15 or 16 or 17 or 18 or 19
21. exp knee/
22. knee*.tw.
23. 21 or 22
24. 20 and 23
25. (rat or rats or pig or pigs or porcine or mice or murine or mouse or sheep or rabbit* or canine or dog*).ti.
26. 24 not 25
27. 14 and 26
28. limit 27 to yr="2004 -Current"

Search strategy for Web of Science Core Collection (2004 to July 2014)

TOPIC: (cost* or economic* or qaly* or "quality of life" or E .More TOPIC: (cost* or economic* or qaly* or "quality of life" or EQ-5D or ICER* or utilit* or health stat* or resource* or SF-36 or short form* or markov or standard gamble or time trade) AND TITLE: (autologous chondrocyte or autologous cartilage or MACI or MACT or chondrocelect) AND TOPIC: (knee*)

Search strategy for NHS Economic Evaluation Database: Issue 2 of 4, April 2014

Search on '(autologous chondrocyte or autologous cartilage or MACI or MACT or chondrocelect) and knee* in Title, Abstract, Keywords in Economic Evaluations'

Appendix 11 Annual transition probabilities

This section reports on the sources of the progression rates used in the Markov model. These transition probabilities were derived from the literature and in consultation with clinical experts.

Most studies presented information in the form of success (progression) rates over a specified time period. These rates were converted to transition probabilities using the formula below, where 'r' is the progression rate and 't' is time:

$$p_{\text{transition}} = 1 - \exp\{-rt\} \quad (8)$$

Where progression rates were not available from the literature, we converted the probability of the event over a period of time to a constant rate using the formula below:

$$r = -[\ln(1-P)]/t \quad (9)$$

Patients receiving an autologous chondrocyte implantation procedure: ACI(ACI), 20–54 years (see Table 57)

Primary repair

Progression rates for people who progressed from 'primary repair' to 'successful primary' and to a 'second repair' were obtained from Saris *et al.*⁷⁰ These authors provided information on a 3-year failure rate of 3.9% for people who required reoperation of the same lesion. The 3-year probability was obtained and then converted to a 1-year transition probability of 0.01317, which was used in the model. They also reported a success rate of 83.0% over a 3-year period. We assumed a 3-year failure rate of 13.1% for people who had no further repair following the primary repair. Three-year probabilities were obtained for these latter two rates and then converted to 1-year transition probabilities.

Successful primary

Progression rates for people who progressed from a successful primary repair to a second repair, and for those who remain in that health state, were based on information from Saris *et al.*¹⁰⁰ These authors reported a response rate of 87.5% over a 2-year time period for people who had undergone a MACI implant. We assumed that 12.5% of the non-responders would move to the 'no further repair' health state and, of these 12.5% patients, we assumed that 10% would move from the 'successful primary repair' to the 'second repair' health state. Based on this information, the following annual transition probabilities were derived: 0.93580, 0.05793 and 0.00627, respectively.

Second repair

The transitions required here include people who have undergone a second repair that was successful, and people in whom it was unsuccessful who have not had a further repair. Saris *et al.*⁷⁰ reported a 3-year success rate for an initial procedure as 83.0%. Here, we assumed that the success rate in the second repair is the same as the success rate in the primary repair, if the second repair is the same as the first. For the people who have no further repair following the second repair, we derived a 1-year transition probability of 0.06022. Here, we assumed that 17.0% of people will have no further repair over a 3-year time period.

Successful second

Progression rates for people who progressed from a 'successful second repair' to 'no further repair', and for those who remain in that health state, were based on information from Saris *et al.*¹⁰⁰ These authors reported a response rate of 87.5% over a 2-year time period for people who had undergone a MACI implant. We assumed that 12.5% of people would move from the 'successful second repair' to the 'no further repair health' state. Based on this information, the following annual transition probabilities were derived: 0.93541 and 0.06459, respectively.

Patients receiving a microfracture procedure: MF(MF), 20–54 years (see Table 57)

Primary repair

People who received a primary repair can remain in the 'successful primary repair' health state, have a second repair, or have no further repair, and these values were obtained from Saris *et al.*⁷⁰ These authors reported that 11.5% of people who had undergone a primary repair required reoperation of the same lesions within 36 months. From this, we derived an annual transition probability of 0.03990 for those who require a second repair. The authors also reported that 62.0% of people will have a successful primary repair within 36 months. Taking account of the 62% who have initial success, and the 11.5% who have a second repair within 3 years, leaves 26.5% of the initial MF group who have no further repair in the first 3 years. We derived an annual transition probability of 0.09754 for people who receive no further repair.

Successful primary

Saris *et al.*¹⁰⁰ reported on the percentage (68.1%) of people who responded to treatment at 2 years. We assumed that 31.9% of the non-responders will move to the 'no further repair' health state and, of these 31.9% patients, we assumed that 10% will move from the 'successful primary repair' to the 'second repair' health state. Based on this information, the following annual transition probabilities were derived: 0.82825, 0.15567 and 0.01608, respectively.

Second repair

Saris *et al.*⁷⁰ reported a 62.0% success rate for people who had an initial primary repair over 36 months. Owing to the paucity of information on the success rate for people receiving a second repair, we assumed the same percentage success for a second repair as for people who had a primary repair. For the people who have no further repair following the second repair, we derived a 1-year transition probability of 0.14730. Here, we assumed that 38.0% of people will have no further repair over a 3-year time period.

Successful second

Saris *et al.*¹⁰⁰ reported on the percentage (68.1%) of people who responded to treatment at 2 years. Here, we assumed that the percentage success for the second repair is the same for people who had a successful primary repair. We assumed 31.9% of people would receive no further repair. From this, we derived an annual transition probability of 0.17477 to represent those people who would receive no further repair. The annual transition probability of 0.82523 was derived to represent people who remained in a 'successful second repair' health state.

Patients receiving microfracture after failed autologous chondrocyte implantation: ACI(MF), 20–54 years (see Table 57)

We report here the values for MF as a second procedure after ACI, as these transition probabilities are different from ACI(ACI).

Second repair

People who had a second repair can have a successful second repair or do not receive a further repair. For those people who do not receive a further repair, we obtained information from Vanlauwe *et al.*⁴³ These authors reported that for 16.4% of people who had the MF procedure following an ACI procedure, this procedure failed at 5 years. From this, we derived an annual transition probability of 0.03519 for people who do not receive a further repair. We assumed the remainder of the people would have a successful MF procedure following an ACI. From this, we derived an annual transition probability from 'second repair' to 'successful second repair' as 0.96481.

Successful second

Saris *et al.*¹⁰⁰ reported that 68.1% of people responded to treatment at 2 years. We assumed that the percentage success for the second MF is as the first MF. We assumed that 31.9% of people would receive no further repair. From this, we derived an annual transition probability of 0.17477 to represent those people who would not receive a further repair. The annual transition probability of 0.82523 was derived to represent people who remained in a 'successful second repair' health state.

Patients receiving autologous chondrocyte implantation after failed microfracture: MF(ACI), 20–54 years (see Table 57)

We report here the values for ACI as a second procedure after MF as these transition probabilities are different from MF(MF).

Second repair

People who had a second repair can have a successful second repair. If the second repair is unsuccessful, we assumed that they do not receive a further repair. For those people who do not receive a further repair, we obtained information from Biant *et al.*⁷⁹ These authors reported that for 30.9% of people who had the ACI procedure following a MF procedure, this procedure failed at 10-year follow-up. From this, we derived an annual transition probability of 0.03629 for people who do not receive a further repair. We assumed the remainder of the people would have a successful ACI procedure following a MF. From this, we derived an annual transition probability from second repair to successful second repair as 0.96371.

Successful second

Saris *et al.*¹⁰⁰ reported that 68.1% of people responded to treatment at 2 years. We assumed that the percentage success for the second repair is the same for people who had a successful primary repair (assuming that this repair was MF). We assumed 31.9% of people would receive no further repair. From this, we derived an annual transition probability of 0.17477 to represent those people who would not receive a further repair. The annual transition probability of 0.82523 was derived to represent people who remained in a 'successful second repair' health state.

Patients 55+ years: all comparisons (see Table 58)

We report here only the transition probability values for the comparisons for patients aged 55+ years, which are different from those for patients aged between 20 and 54 years.

Successful primary, successful second and no further repair

Information required for people who required a TKR was obtained from Knutsen *et al.*⁶⁷ These authors reported that at the 5-year follow-up, of the 40 patients who received an ACI and of the 40 patients who received a MF, nine patients in both groups failed the primary procedure and of these nine patients only one went on to have TKR (the same failure rate for both ACI and MF). For people who require a PKR following a failed primary repair, we assumed that this number would be the same as those receiving a TKR. From this information reported, we derived a 1-year transition probability of 0.00505 to be used in the model for patients moving to the 'first TKR' and 'first PKR' health states from the 'successful primary', 'successful second repair' and 'no further repair' health states.

To estimate values for people who remain in the other health states ('second repair', 'successful second repair' and 'no further repair'), the percentages for TKR and PKR were removed from the totals (i.e. from the success and failure rates) and the annual transition probabilities were re-estimated.

Patients 55+ years: all comparisons (see Table 59)

First total knee replacement

Gerlier *et al.*¹⁴¹ reported information on the percentage success (99%) for people who had a TKR. We assumed this success to be at 5 years following the initial TKR. We derived a transition probability of 0.99223 for patients moving from a first TKR to a successful first TKR. For the progression rates to further knee replacement, Dong and Buxton¹⁵⁸ reported that approximately 2% of people who had undergone their first TKR required a total revision within 2–5 years. Here, we assumed that 2% of people would require a revision procedure in 3.5 years. From this, we derived a 1-year transition probability of 0.00576. We assumed that 1% of people would not receive a further knee replacement 5 years following their first knee replacement.

First partial knee replacement

Owing to the paucity of progression rates available from the literature for people who received a PKR, we used the percentage success and progression for people who received their first TKR. We assumed a transition probability of 0.99223 for a successful first PKR, 0.00576 for people requiring a revision, and 0.00201 for people who receive no further knee replacement.

Successful first total knee replacement

For people who received their primary knee replacement that was successful, we obtained this transition probability from Dong and Buxton.¹⁵⁸ These authors provided information on the 1-month probability of a successful knee replacement and remaining in normal health after the primary TKR. This 1-month probability was converted into a 1-year transition probability of 0.9737. Information on the progression to further knee replacement from a first knee replacement was obtained from Gerlier *et al.*¹⁴¹ These authors reported a 15% revision for people requiring further knee replacement, 15 years after the first TKR. From this, we estimated an annual transition probability of 0.01078 for people requiring further revision. For people who receive no further knee replacement after the initial knee replacement, we derived an annual transition probability based on information on the percentage of successful and revision procedures reported in the studies by Dong and Buxton¹⁵⁸ and Gerlier *et al.*¹⁴¹

Successful partial knee replacement

We assumed the transition probabilities for people who had a PKR to be the same for people who had a TKR. We assumed a 1-year transition probability of 0.97307 for a successful PKR, a probability of 0.01078 for people requiring further revision, and 0.01615 for people who receive no further knee replacement.

Further knee replacement

Gerlier *et al.*¹⁴¹ reported a 90% success rate for people who have received a further knee replacement. We assumed this success to be at 5 years following the further knee replacement. We also assumed that 10% of people would receive no further knee replacement following the further knee replacement. We derived a transition probability of 0.02085 for people requiring no further knee replacements.

Successful further knee replacement

Gerlier *et al.*¹⁴¹ reported a 15% revision rate 15 years after successful TKR. From this we derived a transition probability of 0.01078 for people requiring a further knee replacement. For people who remain in the 'successful further knee replacement' health state following further knee replacement, Gerlier *et al.*¹⁴¹ reported a 90% success rate and we assumed this to be at 5 years. We derived an annual transition probability of 0.97307 for people who remain in this health state. For people who had a successful further knee replacement and require no further knee replacement, we assumed this to be the same as a 1-year transition probability of 0.01615 for successful first TKR and requiring no further knee replacements.

TABLE 57 Annual transition probabilities: 20–54 years

| From\to | Successful primary | Second repair | Successful second | No further repair | Successful primary | Second repair | Successful second | No further repair |
|--------------------|--------------------|---------------|-------------------|-------------------|--------------------|---------------|-------------------|-------------------|
| | ACI(ACI) | | | | MF(MF) | | | |
| Primary repair | 0.94110 | 0.01317 | – | 0.04573 | 0.86256 | 0.03990 | – | 0.09754 |
| Successful primary | 0.93580 | 0.00627 | – | 0.05793 | 0.82825 | 0.01608 | – | 0.15567 |
| Second repair | – | – | 0.93978 | 0.06022 | – | – | 0.85270 | 0.14730 |
| Successful second | – | – | 0.93541 | 0.06459 | – | – | 0.82523 | 0.17477 |
| No further repair | – | – | – | 1.00000 | – | – | – | 1.00000 |
| | MF(ACI) | | | | ACI(MF) | | | |
| Primary repair | 0.86256 | 0.03990 | – | 0.09754 | 0.94110 | 0.01317 | – | 0.04573 |
| Successful primary | 0.82825 | 0.01608 | – | 0.15567 | 0.93580 | 0.00627 | – | 0.05793 |
| Second repair | – | – | 0.96371 | 0.03629 | – | – | 0.96481 | 0.03519 |
| Successful second | – | – | 0.82523 | 0.17477 | – | – | 0.82523 | 0.17477 |
| No further repair | – | – | – | 1.00000 | – | – | – | 1.00000 |

TABLE 58 Annual transition probabilities: 55+ years

| From\to | Successful primary | Second repair | Successful second | No further repair | First TKR | First PKR | Successful primary | Second repair | Successful second | No further repair | First TKR | First PKR |
|--------------------|--------------------|---------------|-------------------|-------------------|-----------|-----------|--------------------|---------------|-------------------|-------------------|-----------|-----------|
| | ACI(ACI) | | | | | | | | | | | |
| Successful primary | 0.95180 | 0.00376 | - | 0.03434 | 0.00505 | 0.00505 | 0.84726 | 0.01608 | - | 0.12656 | 0.00505 | 0.00505 |
| Second repair | - | - | 0.93978 | 0.06022 | - | - | - | - | 0.85270 | 0.14730 | - | - |
| Successful second | - | - | 0.95167 | 0.03823 | 0.00505 | 0.00505 | - | - | 0.84489 | 0.14501 | 0.00505 | 0.00505 |
| No further repair | - | - | - | 0.98990 | 0.00505 | 0.00505 | - | - | - | 0.98990 | 0.00505 | 0.00505 |
| | MF(ACI) | | | | | | | | | | | |
| Successful primary | 0.84726 | 0.01608 | - | 0.12656 | 0.00505 | 0.00505 | 0.95180 | 0.00376 | - | 0.03434 | 0.00505 | 0.00505 |
| Second repair | - | - | 0.96371 | 0.03629 | - | - | - | - | 0.96481 | 0.03519 | - | - |
| Successful second | - | - | 0.84489 | 0.14501 | 0.00505 | 0.00505 | - | - | 0.84489 | 0.14501 | 0.00505 | 0.00505 |
| No further repair | - | - | - | 0.98990 | 0.00505 | 0.00505 | - | - | - | 0.98990 | 0.00505 | 0.00505 |
| | MF(MF) | | | | | | | | | | | |
| Successful primary | 0.84726 | 0.01608 | - | 0.12656 | 0.00505 | 0.00505 | 0.95180 | 0.00376 | - | 0.03434 | 0.00505 | 0.00505 |
| Second repair | - | - | 0.96371 | 0.03629 | - | - | - | - | 0.96481 | 0.03519 | - | - |
| Successful second | - | - | 0.84489 | 0.14501 | 0.00505 | 0.00505 | - | - | 0.84489 | 0.14501 | 0.00505 | 0.00505 |
| No further repair | - | - | - | 0.98990 | 0.00505 | 0.00505 | - | - | - | 0.98990 | 0.00505 | 0.00505 |
| | ACI(MF) | | | | | | | | | | | |
| Successful primary | 0.84726 | 0.01608 | - | 0.12656 | 0.00505 | 0.00505 | 0.95180 | 0.00376 | - | 0.03434 | 0.00505 | 0.00505 |
| Second repair | - | - | 0.96371 | 0.03629 | - | - | - | - | 0.96481 | 0.03519 | - | - |
| Successful second | - | - | 0.84489 | 0.14501 | 0.00505 | 0.00505 | - | - | 0.84489 | 0.14501 | 0.00505 | 0.00505 |
| No further repair | - | - | - | 0.98990 | 0.00505 | 0.00505 | - | - | - | 0.98990 | 0.00505 | 0.00505 |

TABLE 59 Annual transition probabilities: 55+ years (for all scenarios)

| From\to | Successful first TKR | Successful first PKR | Further knee replacement | Successful further knee replacement | No further knee replacement |
|-------------------------------------|----------------------|----------------------|--------------------------|-------------------------------------|-----------------------------|
| All comparisons | | | | | |
| First TKR | 0.99223 | – | 0.00576 | – | 0.00201 |
| First PKR | – | 0.99223 | 0.00576 | – | 0.00201 |
| Successful first TKR | 0.97307 | – | 0.01078 | – | 0.01615 |
| Successful first PKR | – | 0.97307 | 0.01078 | – | 0.01615 |
| Further knee replacement | – | – | – | 0.97915 | 0.02085 |
| Successful further knee replacement | – | – | 0.01078 | 0.97307 | 0.01615 |
| No further knee replacement | – | – | – | – | 1.00000 |

TABLE 60 Sensitivity analysis: mean utility values used in the economic model

| Repairs | First repair | | | | Source |
|--|--------------|-------|-------|-------|--|
| | ACI | MF | ACI | MF | |
| Before primary repair | 0.519 | | | | Oswestry submission ³⁵ |
| Successful primary | | | | | |
| First year | 0.645 | 0.627 | | | Oswestry submission ³⁵ |
| Second year | 0.669 | 0.667 | | | |
| Third year | 0.674 | 0.673 | | | |
| Fourth year | 0.647 | 0.674 | | | |
| Five years plus | 0.698 | 0.658 | | | |
| Before second repair | 0.519 | | | | |
| Choose not to have a second repair | 0.600 | | | | |
| Second repair | | | | | |
| Successful second | | | | | |
| First year | 0.645 | 0.627 | 0.645 | 0.627 | Oswestry submission ³⁵ |
| Second year | 0.669 | 0.667 | 0.669 | 0.667 | |
| Third year | 0.674 | 0.673 | 0.674 | 0.673 | |
| Fourth year | 0.647 | 0.674 | 0.647 | 0.674 | |
| Five years plus | 0.698 | 0.658 | 0.698 | 0.658 | |
| No further repair | 0.600 | | | | |
| Replacements | | | | | |
| Before first knee replacement (TKR) | 0.615 | | | | Dong and Buxton, ¹⁵⁸ Jansson and Granath ¹⁵⁹ |
| Before first knee replacement (PKR) | 0.615 | | | | |
| Successful first knee replacement: TKR | 0.780 | | | | Dong and Buxton ¹⁵⁸ |
| Successful first knee replacement: PKR | 0.780 | | | | |
| Before further TKR | 0.557 | | | | Gerlier <i>et al.</i> ¹⁴¹ |
| Successful further TKR | 0.780 | | | | Dong and Buxton ¹⁵⁸ |
| No further TKR | 0.691 | | | | Gerlier <i>et al.</i> ¹⁴¹ |

Some patients decide not to have another repair attempt after unsuccessful first repair. We have assumed that they had the same benefit and do not go back to their baseline utility.

A decorative graphic consisting of numerous thin, parallel green lines that curve from the left side of the page towards the right, creating a sense of movement and depth.

**EME
HS&DR
HTA
PGfAR
PHR**

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

This report presents independent research funded by the National Institute for Health Research (NIHR). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health

Published by the NIHR Journals Library