Reviewing Research Evidence for Nursing Practice: Systematic Reviews

Edited by

Christine Webb

Professor of Health Studies
Faculty of Health and Social Care, University of Plymouth, UK

and

Brenda Roe

Professor of Health Sciences
Institute of Health Research, Faculty of Health and Applied Social Sciences, Liverpool John Moores University, UK

Blackwell
Publishing

Reviewing Research Evidence for Nursing Practice

Reviewing Research Evidence for Nursing Practice: Systematic Reviews

Edited by

Christine Webb

Professor of Health Studies
Faculty of Health and Social Care, University of Plymouth, UK

and

Brenda Roe

Professor of Health Sciences
Institute of Health Research, Faculty of Health and Applied Social Sciences, Liverpool John Moores University, UK

Blackwell
Publishing

Blackwell Publishing editorial offices:
Blackwell Publishing Ltd, 9600 Garsington Road, Oxford OX4 2DQ, UK
Tel: +44 (0)1865 776868
Blackwell Publishing Inc., 350 Main Street, Malden, MA 02148-5020, USA
Tel: +1 781 388 8250
Blackwell Publishing Asia Pty Ltd, 550 Swanston Street, Carlton, Victoria 3053, Australia
Tel: +61 (0)3 8359 1011

First published 2007 by Blackwell Publishing Ltd

ISBN: 978-1-4051-4423-0

Library of Congress Cataloging-in-Publication Data
Reviewing research evidence for nursing practice : systematic reviews / edited by Christine Webb and Brenda Roe.
p. ; cm.
Includes bibliographical references and index.
ISBN-13: 978-1-4051-4423-0 (pbk. : alk. paper)
1. Nursing—Research—Methodology. 2. Systematic reviews (Medical research) 3. Evidence-based nursing. I. Webb, Christine. II. Roe, Brenda H.
[DNLM: 1. Clinical Nursing Research. 2. Review Literature. 3. Meta-Analysis. WY 20.5 R454 2007]
RT81.5.R488 2007
610.73072—dc22
2007010023

A catalogue record for this title is available from the British Library

Set in 9.5/11.5pt Palatino
by Graphicraft Limited, Hong Kong
Printed and bound in Singapore
by Fabulous Printers Pte Ltd

For further information on Blackwell Publishing, visit our website:
www.blackwellnursing.com

Contents

Contributors

Donna Ciliska is Professor in the School of Nursing at McMaster University and has an appointment as a nursing consultant with Hamilton Public Health. She is editor of the journal *Evidence-Based Nursing*, and has contributed as a co-editor to two evidence-based nursing texts. Her research interests include community health, obesity, eating disorders and knowledge translation.

Mike Clarke is Director of the UK Cochrane Centre, which provides training and support to systematic reviewers in the UK. He is Professor of Clinical Epidemiology at the University of Oxford, working on systematic reviews of individual patient data. These include the breast cancer overview, which brings together data on more than 300 000 women in 400 randomised trials. He works on more than a dozen other systematic reviews across health care and on trials in pre-eclampsia, subarachnoid haemorrhage, breast cancer and poisoning – which are the world's largest randomised trials in each condition.

Maureen Dobbins is an associate professor in the School of Nursing at McMaster University and has an appointment as a nursing consultant with the City of Hamilton Public Health Services. She holds a career scientist award with the Ontario Ministry of Health and Long-Term Care. Her research interests include knowledge transfer and exchange, evidence-informed decision-making, community health, healthy body weight, physical activity and chronic disease prevention.

Chantale L. Dumoulin is an assistant professor at the School of Physical and Occupational Therapy, Faculty of Medicine, University of Montreal, Canada. Her research interests include psychometric evaluation of measuring instruments, conservative interventions for urinary incontinence in women, service delivery and research dissemination.

David Evans is Senior Lecturer in the Division of Health Sciences at the School of Nursing and Midwifery, University of South Australia. His areas of interest include all aspects of acute care nursing, evidence-based practice, safety and quality issues and practice evaluation.

Beverley French is a senior research fellow at the University of Central Lancashire. Her experience of quantitative systematic review is mainly in Cochrane reviews of interventions in stroke rehabilitation. She is currently involved in a number of syntheses of wider evidence sources relating to mental health advocacy, and community development and engagement.

E. Jean C. Hay-Smith is a lecturer in the Rehabilitation Teaching and Research Unit at the Wellington School of Medicine and Health Sciences, University of Otago, New Zealand. She is an editor of the Cochrane Incontinence Review Group. Her research interests include the conservative management of urinary incontinence (particularly pelvic floor muscle training), self-efficacy and treatment adherence.

Peter Herbison works in the Department of Preventive and Social Medicine at the University of Otago in Dunedin, New Zealand, providing statistical help for researchers.

Myfanwy Lloyd Jones is a senior research fellow in the Health Economics and Decision Science Section of the University of Sheffield School of Health and Related Research (ScHARR). A specialist in systematic reviewing, she is a member of the ScHARR Technology Assessment Group (ScHARR-TAG), and has contributed to a number of technology assessments within the NHS Health Technology Assessment Programme.

Rachel McNamara is a research fellow in the Department of General Practice, Cardiff University, UK.

Lynn Nicholls is Lecturer in Midwifery at the University of Plymouth in Taunton, UK.

Christine Norton is Nurse Consultant (Bowel Continence) at St Mark's Hospital in Harrow and Burdett Professor of Gastrointestinal Nursing, King's College London, UK. She is an editor for the Cochrane incontinence group, chairs the Royal College of Nursing Gastroenterology and Stoma Care Forum and is associate editor of *Gastrointestinal Nursing*.

Beverly O'Connell holds the Inaugural Chair in Nursing at the Deakin-Southern Health Nursing Research Centre. Previously she held the positions of Chair in Nursing, Cabrini Health; Nursing Research Director, Sir Charles Gairdner Hospital; and Director of Nursing Research and Development, Curtin University, Australia. Her research interests include aged care, quality and safety, incontinence care and carer support.

Joan Ostaszkiewicz is a research fellow and PhD candidate at the School of Nursing at Deakin University. She holds a joint appointment with the Southern Health Network and Deakin University. Her research interests are ageing and the integration of research with practice.

Barbara L. Paterson holds a Tier 1 Canada Research Chair in Chronic Illness and is a professor at the University of New Brunswick in Canada. She is widely published in the fields of chronic illness and qualitative research.

Brenda Roe is Professor of Health Sciences at Liverpool John Moores University. She has a background in nursing, health visiting, primary care, public health, gerontology, health services research and management. She is a Fellow of the Queen's Nursing Institute and Fellow of the Royal Society for the Promotion of Health.

Margarete Sandelowski is Cary C. Boshamer Professor in the School of Nursing at the University of North Carolina at Chapel Hill, USA. She is Director of the Annual Summer Institutes in Qualitative Research, and of the new Certificate Program in Qualitative Research, both offered at the School of Nursing. She has published widely in nursing and social science anthologies and journals in the areas of technology and gender, especially reproductive technology and technology in nursing, and of qualitative methods. Her latest book, co-authored with Julie Barroso, is *Handbook for Synthesizing Qualitative Research* (Springer, New York, 2007).

Chris Shaw is Reader in Nursing Research in the School of Care Sciences at the University of Glamorgan, South Wales, UK. Her research interests focus on chronic disease management and health behaviours such as self-care and help-seeking. She has a background in nursing and midwifery and is a chartered health psychologist.

Helen Thomas is an Associate Professor in the School of Nursing, McMaster University and a Clinical Consultant with the Public Health Research, Education and Development Program, Hamilton, Ontario, Canada, where she is Project Leader of the Effective Public Health Practice Project.

Lois Thomas is Senior Research Fellow in the Department of Nursing at the University of Central Lancashire. Her research interests include stroke nursing, particularly urinary incontinence after stroke, and the effectiveness of clinical guidelines in nursing and allied health professions.

Christine Webb is Professor of Health Studies at the University of Plymouth, UK, Executive Editor of *Journal of Advanced Nursing*, and Editor of *Nurse Author & Editor*. Her initial clinical specialism was women's health, but more recently she has focused on nurse education as a manager and researcher. She is a Fellow of the Royal College of Nursing (UK).

Robin Whittemore is Associate Professor at the Yale School of Nursing in New Haven CT, USA. Her research interests include lifestyle change, nurse-coaching, type 2 diabetes, psychosocial adjustment to chronic illness, nursing intervention research, and nursing theory development.

Preface

We know from many research studies that practising nurses and other healthcare professionals do not always have the time, confidence or skills to carry out research or systematic reviews for themselves. Therefore they rely on reviews by other people when considering innovations and developments in their practice.

Our aim for this book, therefore, is to present readers with the issues arising from conducting systematic reviews and thereby to help them understand reviews that they identify and read when considering developing their health policy, services and clinical practice.

It is not solely a 'how to do a systematic review' book – as other examples of that have already been published. Rather, we have presented how a selection of reviews has been carried out in a range of specialist areas related to health policy, service development and clinical practice. This will help readers to critically appraise the reviews they read and judge how useful they are for changing practice and service development. A particular novel and groundbreaking feature of this book is that it includes examples of all types of review – quantitative, qualitative and integrative or mixed-method reviews which include both qualitative and quantitative empirical studies – whereas other books are limited to only one of these types. By bringing all these approaches together in one book, we hope to offer a reader-friendly and economical volume for nurses, healthcare professionals and health-services researchers.

The book will be of interest to nurses and healthcare professionals in practice, people following an MSc or taught doctorate programme in advanced or specialist practice or postgraduate study, as well as academic researchers and research doctorate students.

Introduction

Brenda Roe and Christine Webb

From the early 1990s systematic review as a method of establishing the evidence of effectiveness of healthcare interventions has developed apace – most notably, with the development of the international Cochrane Collaboration and the Cochrane Library for the electronic dissemination of systematic reviews. These reviews focus on quantitative evidence from randomised controlled trials and meta-analyses. Parallel developments, but not on the large international scale of the Cochrane Collaboration, have also evolved looking at the meta-study and meta-synthesis of qualitative research evidence. Methods, handbooks, critical appraisal and quality criteria are available and are described in this book. More recently, integrative reviews are being developed to combine the evidence from quantitative research and qualitative research on clinical topics, management and policy, as undertaken by the Joanna Briggs Foundation. It is acknowledged that the methodology and methods for systematic reviews are developing and increasingly need to take account of diverse sources of evidence (Popay, 2006), along with the recognition and development of terms and definitions (Sander & Kitcher, 2006).

The purpose of this book is to present the issues arising when conducting systematic reviews and to provide a 'how to' of the methods used, based on reviewers' experiences of undertaking published systematic reviews. It provides a selection of reviews carried out in a range of specialist areas related to clinical practice, along with recommendations for practice and future research. Not only does the book inform people wishing to undertake systematic reviews themselves, but also clinicians who may wish to appraise the reviews they read with a view to incorporating their recommendations into practice. It is known from many research studies that practising clinicians do not have the time, confidence or skills to carry out research and they rely on reviews undertaken by others when considering innovations and developments in their clinical practice.

The book is novel and is the only one of its kind to include systematic reviews of quantitative research, qualitative research, and integrative reviews incorporating both quantitative evidence and qualitative evidence. The methods for systematic reviews are continuing to evolve and this book provides an indication of this evolution in one volume. The book is primarily intended for nurses and nursing, but is of relevance to medical and health services researchers and clinicians as well as those from the professions allied to medicine.

The book is in four parts. Part 1 covers Systematic Reviews and Meta-Analysis of Quantitative Research and predominantly cites as examples reviews undertaken as part of the Cochrane Collaboration involving randomised controlled trials. Part 2, entitled Meta-synthesis and Meta-study of

Qualitative Research, includes systematic reviews of qualitative evidence and studies, while Part 3 includes Integrative Reviews of Quantitative and Qualitative Research. Finally, Part 4 looks at the Application and Uses of Reviews in health services as well as offering reflections on the past, present and future of systematic reviews.

Each of the chapters begins with an Introduction to set the clinical context and concludes with implications for practice and future research. In Part 1, Chapter 1, an Overview of Methods by Mike Clarke, gives an overview of systematic review methods for quantitative studies, notably randomised controlled trials, and includes methods for locating, appraising and combining independent studies that are transparent and minimise bias. Such reviews place research in context and ensure that new research is developed and implemented appropriately. Systematic reviews are increasingly more common, as exemplified by the endeavours of the Cochrane Collaboration and the Cochrane Library based on a global effort established in 1993. Clarke's chapter looks at question formulation, study identification, appraisal of studies for inclusion, data collection, statistical analysis, updating of reviews and appraising and using systematic reviews. He concludes that systematic reviews offer the best way to ensure that evidence is available on which to make decisions.

Chapter 2 is by Brenda Roe and includes Key Stages and Considerations when Undertaking a Systematic Review. The Cochrane systematic review on bladder training for the management of urinary incontinence in adults is used as an example and sections of the chapter include guidelines, developing a protocol and necessary steps, literature searching, publication bias, inclusion and exclusion criteria, quality assessment, data extraction, outcomes, review methods, presentation, and combining and interpretation of results, along with statistical outcome measures and combined effect estimates. The chapter is supported with figures and tables as examples that can be used by people wishing to undertake future systematic reviews, and concludes with sections on writing up and disseminating reviews.

Chapter 3, entitled Prevention and Treatment of Urinary Incontinence After Stroke in Adults: Experiences, is based on a systematic review for the Cochrane Collaboration by Lois Thomas and Beverley French. It provides an overview of the methods used and the reviewers' conclusions, followed by sections on issues that arose when carrying out the review, designing the protocol, designing the search, retrieval of potential studies for inclusion, data extraction and assessment of study quality. Sections on extraction of outcome data, data analysis and synthesis are followed by valuable learning points which are of direct benefit for people wishing to undertake future systematic reviews. The chapter concludes not only with implications for practice but also with lessons for future similar reviews.

Chapter 4, like Chapters 2 and 3, also focuses on a Cochrane systematic review on urinary incontinence as an example. It is entitled Pelvic Floor Muscle Training for Urinary Incontinence in Women and is by Jean Hay-Smith, Chantale Dumoulin and Peter Herbison. An overview of the review is provided, along with conventional subject headings followed by a discussion and the issues that arose when carrying out the review. These include sections on methodological heterogeneity, other sources of heterogeneity, and choice and reporting of outcome measures. Their chapter illustrates the evolving nature and complexity of randomised controlled trials designs and methods.

Chapter 5, the last chapter in Part 1, is by Christine Norton and also includes a Cochrane systematic review by way of example, entitled Biofeedback and Anal Sphincter Exercises for Faecal Incontinence in Adults. Faecal incontinence, biofeedback and exercises are set in context, followed by an overview of the review and its methods, results and conclusions. Issues that arose while carrying out the review included randomised versus non-randomised evidence, outcome measures, international relevance and translating the evidence into clinical recommendations. The chapter concludes by discussing the relationship of the review with other systematic reviews on the subject and with reflections for future reviews.

Part 2 is a section on Meta-study and Meta-synthesis of Qualitative Research, with Myfanwy Lloyd Jones in Chapter 6 including an Overview of the Methods in which both meta-study and meta-synthesis are defined. She provides a brief history and then goes on to cover key methodological aspects, such as the focus of the study, inclusion and exclusion criteria and theoretical framework.

This is followed by sections on study identification and selection, summary, analysis and synthesis of findings. The chapter is completed by presenting the interpretation of results and dissemination of findings, along with assessing the quality of meta-syntheses.

Chapter 7 looks at Coming Out as Ill: Understanding Self-Disclosure in Chronic Illness from a meta-synthesis of qualitative research by Barbara L. Paterson. The chapter includes primary research and deals with sample characteristics, preparing for the meta-study, analytic components, meta-synthesis, challenges in meta-study projects, conducting a meta-study alone and issues of selecting the primary research to be included.

Chapter 8 is entitled From Meta-synthesis to Method: Appraising the Qualitative Research Synthesis Report and is written by Margarete Sandelowski. She looks at the components of the qualitative research synthesis report and evaluation criteria and methods, using her study of prenatal diagnosis as an example. Qualitative research synthesis is contrasted with narrative overview, synthesis of quantitative research findings, secondary analysis, within-study and within-programme research synthesis and meta-study, and this is followed by consideration of results and discussion of the synthesis produced.

Chapter 9 completes Part 2 and is by Myfanwy Lloyd Jones, who presents her study on Role Development in Acute Hospital Settings: A Systematic Review and Meta-synthesis. She gives an overview of the methods used and aim of the study, which looked at innovative roles of nurses, and barriers and facilitators, and used Paterson's meta-study methodology (see Chapter 7). Conventional section headings of methods, results and findings are included, followed by discussion of issues that arose while carrying out the review, identifying potentially relevant studies and retrieving them, data extraction and study appraisal. Sections on meta-data-analysis and meta-synthesis follow, and the chapter concludes with consideration of interpretation of the results and limitations.

Part 3 is particularly novel and covers Integrative Reviews of Quantitative and Qualitative Research. Chapter 10 by David Evans provides an Overview of Methods and looks at rigour in integrative reviews, systematic methods, problem and purpose, literature searching and data collection. He continues with sections on evaluation of the quality of primary research, evidence of critical appraisal, and transparency, and concludes by considering quality in integrative reviews.

Chapter 11, entitled Rigour in Integrative Reviews, by Robin Whittemore develops some of these themes. She starts by considering what are integrative reviews, their purpose, the review protocol, problem identification and location of studies. She provides details about evaluating studies, data collection and analysis – specifically descriptive data synthesis, statistical data synthesis and qualitative data synthesis – along with a section on the integrative review report.

Chapter 12 is by Joan Ostaszkiewicz and Beverly O'Connell and looks at Habit Retraining for Urinary Incontinence in Adults. It builds on a Cochrane systematic review of quantitative evidence from randomised controlled trials and synthesises evidence from other study designs to provide an integrative review on the topic. As well as conventional method sections and related considerations, they include discussion of the dilemmas they encountered in implementing the Cochrane systematic review criteria, in limiting the review to one form of evidence, as well as with critical appraisal and establishing levels of quality. They go on to detail managing and integrating evidence from mixed design studies, using habit retraining as the example.

Chapter 13 addresses the question What Makes a Good Midwife? and is by Lynn Nicholls and Christine Webb, who undertook an integrative review to answer this question. They give an overview of the methods, protocol and search methods, appraisal of studies, analysis of findings as well as discussing methodological issues. The chapter is completed with a summary of the main findings, aspects of conducting an integrative review and issues that arose.

In Chapter 14, Rachel McNamara and Chris Shaw present an integrative review investigating Older People and Respite Care. They address the questions of who are carers and what impact their role has on them, and then go on to consider respite care and evidence of its effectiveness. They provide an overview of the research aims, methodology and methodological issues. They consider how to devise an appropriate search strategy to capture both quantitative and qualitative evidence, along

with assessment of study quality – which for quantitative studies is more established than for qualitative studies (see chapters in Parts 1 and 2). The identification of studies, data extraction, analysis plan and data synthesis are considered, along with lessons for future reviews.

Part 3 concludes with Chapter 15 by David Evans, which presents an integrative review on the Use of Physical Restraint. As well as methodological considerations, he provides a synthesis of results and lessons learned on use of physical restraint, characteristics of restrained people, reasons for restraining people, injury and physical restraint, the experience of physical restraint and restraint minimisation.

Finally, Part 4 considers the Applications and Uses of Reviews, with Chapter 16 providing steps, methods and considerations for Using Systematic Reviews in Health Services; this chapter is written by Donna Ciliska, Maureen Dobbins and Helen Thomas. They look at how systematic reviews have been used to inform clinical practice, management and policy development by critically appraising reviews using explanation and application of criteria to existing systematic reviews and clinical scenarios with a public health and health promotion focus. The clinical scenarios include teenage suicides and type 2 diabetes mellitus, and include sections on finding the evidence and critical appraisal.

Chapter 17 by Christine Webb and Brenda Roe concludes the volume by summarising the chapters and offering Reflections on the Past, Present and Future of Systematic Reviews. It sets systematic reviews in historical context, from the evolution of systematic reviews for quantitative evidence, then the synthesis of qualitative evidence, followed more recently by integrative reviews which combine analysis and synthesis of both types of evidence in a review. Finally, the possibility of undertaking synopses of a number of related systematic reviews using meta-study techniques is suggested.

Contributors are drawn from a variety of professional disciplines and countries around the globe, reflecting the interdisciplinary nature of systematic reviewing and the international collaborations and networks that have been formed. We are indebted to and would like to thank our contributing authors, who are not only pioneers in their fields but generous individuals willing to communicate effectively and share their expertise with the wider community, despite having busy schedules and workloads.

References

Popay, J. (ed.) (2006) *Moving Beyond Effectiveness in Evidence Synthesis: Methodological Issues in the Synthesis of Diverse Sources of Evidence*. National Institute for Health and Clinical Excellence, London. (NICE website www.nice.org.uk)

Sander, L. & Kitcher, H. (2006) *Systematic and Other Reviews: Terms and Definitions Used by UK Organisations and Selected Databases. Systematic Review and Delphi Survey*. National Institute for Health and Clinical Excellence, London. (NICE website www.nice.org.uk)

Part 1

Systematic Reviews and Meta-Analysis of Quantitative Research

1 Overview of Methods

Mike Clarke

Introduction

Systematic reviews are both scientific research and the application of common sense. They serve to identify studies relevant to a particular question, to appraise and assess the eligibility of these studies, and to summarise them, using statistical techniques to combine their results, if feasible and appropriate. Without systematic reviews, we are faced with an ever-increasing number of individual studies. There may be many, sometimes hundreds, on the same question. If this research is to be used to make well-informed decisions, we need to be confident that the effects of both bias and chance are minimised. These effects must be minimised not only within the individual studies but also in the process of bringing them together in a review.

This is where systematic reviews are especially helpful. Regardless of whether the underlying research comprises randomised trials assessing the relative effects of different interventions, studies of test accuracy to determine which is the best technique to diagnose an illness, cohort studies to estimate the prognosis of patients with different characteristics, or qualitative research to understand better the ways in which people make choices, systematic reviews of the research appropriate to answer a question will provide someone making decisions with a more reliable basis for doing so than an individual study.

Systematic reviews are pieces of research, which aim to identify, appraise and summarise studies of relevance to a particular topic. Such a review uses a predefined, explicit methodology, setting out the objectives, eligibility criteria and methods for the review. These methods should be chosen so as to minimise bias in all aspects of the conduct and reporting of the review; including study identification, assessment of eligibility, collection of data, analyses and interpretation. A systematic review does not need to combine the results of the studies to provide an average estimate but, if it does so, this should also be done in a way that minimises bias, with a clear separation between hypothesis testing and hypothesis generating results. This chapter outlines some of these key features of systematic reviews, setting the scene for the more detailed discussion and examples that follow.

Background

Most individual pieces of research are too small on their own to answer reliably all the questions addressed by the research or of relevance to a person wishing to use the research when making a decision about health care. Individual studies may be subject to biases in regard to their availability and might not contain a sufficiently large number or range of participants. Chance effects may lead to

an overestimate or underestimate of the true effect in any scientific investigation. For example, even the best-conducted randomised trial is not immune to the effects of chance and there is no way of knowing whether chance has caused its result to be better or worse than it should be. To minimise the effects of chance, the results of similar studies can be combined – in a meta-analysis – to produce a statistically more reliable result. To minimise the effects of bias, as many as possible of the eligible studies need to be identified and their quality and relevance need to be assessed.

The narrative review article has long been a feature of the healthcare literature, but systematic reviews represent an important departure from these. In a systematic review, the methods used to locate, appraise and, where appropriate, combine independent studies are clearly described. These methods should be transparent and should minimise the possibility of bias.

Systematic reviews are needed both to place research in context and also to ensure that new research is designed and implemented in the most appropriate way (Clarke, 2004). They are increasingly common, not least through the work of The Cochrane Collaboration. This global effort was established in 1993 (Chalmers, 1993) and more than 14 000 people in 100 countries are now involved in its efforts to prepare, maintain and promote the accessibility of systematic reviews of the effects of healthcare interventions (www.cochrane.org). Through this work, the *Cochrane Database of Systematic Reviews* (*CDSR*) now contains the full text for more than 3000 Cochrane systematic reviews, with protocols for 1600 more that are in progress also published in *CDSR*, which is available in *The Cochrane Library* (www.thecochranelibrary.com). There are also several thousand other systematic reviews of the effects of healthcare interventions in the literature; as well as a small, but growing, number of systematic reviews of other aspects of health.

Question formulation and study identification for a systematic review

A systematic review would usually aim to identify and include all research relevant to the question for the review. This objective might be driven by a desire to provide as precise an estimate as possible of the relative effects of two treatments. But it might also be driven by a desire to bring together as much relevant research as possible so as to describe what has already been done, to help ensure that new research learns from the successes and failures of the past, and to identify gaps in the research base (Alderson & Roberts, 2000). Whichever type of review is to be done, the most important first step is the same as that for any research – decide upon the objectives and the questions to be tackled by the systematic review. This will have an impact on the inclusion and exclusion criteria for the review. These might be set out by describing the types of study design, participants, interventions and outcome measures that would be relevant.

When this has been decided, the systematic process for identifying relevant studies begins. Collecting all studies – irrespective of their results – will remove any biases that would be introduced if research with positive results, or which agrees most closely with the opinions and prejudices of the person doing the review, was sought preferentially over other research. Finding and using the results of all relevant studies will minimise chance effects by maximising the amount of data available for analysis and, hence, improve the precision of the estimate in the meta-analyses.

The ideal systematic review is one in which all the relevant studies have been identified before their results could influence decisions about their inclusion. This would overcome the problem of publication bias and of other biases where prior knowledge of the results of a study might influence the reviewer's decision on whether it should be included in her review. However, it needs to be remembered that systematic reviews are, by their nature, a form of retrospective research. The reviewers might already know of some of the potentially eligible studies, and their results. If the systematic review is transparent about the choices made when it was done and strived to find studies beyond those that were already known to the reviewer, users of their review can be more confident that its conduct was not overly influenced or biased by this prior knowledge.

The problem of publication bias makes the search for relevant studies especially difficult, and it will only be overcome through initiatives such as prospective registration of studies at inception (Dickersin et al., 1992). Publication bias usually

arises because studies are more likely to be written up and published if they have statistically significant positive results. A more general rule is that whether or not a trial is published might be influenced by its results. This means that the results of published and unpublished trials might be systematically different. Therefore, unless all trials are sought regardless of their publication status, the systematic review may contain a biased set of studies. In such a case, regardless of the data collection and statistical methods used, a meta-analysis based on these studies may be mathematically precise but clinically wrong. Therefore, unpublished research and studies published only as abstracts or in journals that are difficult to obtain must be sought. This may require extensive searching of relevant bibliographic databases and of journals and conference proceedings (Hopewell et al., 2002), with attention also being given to strategies to find studies published in languages other than English (Pilkington et al., 2005).

The ease of finding randomised trials for systematic reviews has increased throughout the past decade. This is largely through the work of members of The Cochrane Collaboration who have hand searched journals and conference abstracts from cover to cover, looking for reports of randomised trials, and have conducted extensive electronic searching of bibliographic databases. In 1993, fewer than 20 000 reports of randomised trials could be found easily in MEDLINE, even though that database alone contained several tens of thousands more such reports. The Cochrane Collaboration's efforts to identify and make accessible information on reports of trials that might be suitable for inclusion in Cochrane reviews have led to the re-indexing of many of these reports in MEDLINE. Furthermore, the Collaboration, with coordination by the US Cochrane Center, built the *Cochrane Central Register of Controlled Trials* (*CENTRAL*) as a repository of records relating to controlled trials. These include records from MEDLINE and EMBASE and also tens of thousands of records that are in neither database. *CENTRAL* is, therefore, a unique resource for reviewers searching for randomised trials (Dickersin et al., 2002). Unfortunately, reviewers for whom other types of study would be eligible for their review are not so fortunate and still need to rely on their own extensive searches of databases, journals, conference proceedings, etc.

Appraising studies for inclusion in a systematic review

Assessing the eligibility of studies for a systematic review is a key step in determining that the studies meet the inclusion criteria and are of appropriate quality. Many tools are available for assessing the quality of randomised trials but caution is needed in using these. As Juni and colleagues have shown, different quality instruments can give widely different findings (Juni et al., 1999). Rather, it may be preferable for the reviewers to decide upon the key aspects of quality for studies in their review and then to appraise and describe each study on this basis. In randomised trials, these aspects might relate to the generation and concealment of the randomisation schedule, blinding or masking of the interventions, and loss to follow-up. Tools and means to assess the quality of non-randomised trials have also been developed, and some of these have been identified as particularly suitable for use in systematic reviews (Deeks et al., 2003). The distinction between being able to assess the quality of a report, rather than the quality of the underlying study, also needs to be kept in mind (Soares et al., 2004).

Whichever technique is used to assess the quality of the studies in the review, reviewers should also consider how they will use their conclusions about study quality in their review (Detsky et al., 1992). For example, if a systematic review is designed to generate as reliable an estimate as possible of the effects of an intervention, poor-quality studies might be excluded from this calculation. Whereas, if the review seeks to map out what is good and bad about prior research, the inclusion of poor-quality studies would add to the richness of this discussion.

Collection of data

Having decided on the studies that are eligible for the review, the reviewer then needs to gather together information and data on these studies. Even if there is no intention to do a meta-analysis, this information will help to highlight differences and similarities between the studies and will also make it easier to summarise each study and its findings in a standardised way. This should make it easier for the user of the review to compare and

contrast these studies. The reviewer needs to decide how much or how little information to extract for each study, and what sources will be used if the published reports contain insufficient information (Clarke & Stewart, 1994). In compiling as complete a dataset as feasible and sensible, the principles of minimising systematic biases and chance effects must be applied. All relevant trials should be included in the meta-analysis and, if this is not possible, any trials that do not contribute data must not be so numerous or atypical that they introduce important bias to the result of the meta-analysis. If the results of a study have not been published or have only been published in part, the reviewer will need to contact the researchers responsible to try to obtain the necessary data. This can take time and there is no guarantee of success. However, without these data, there is a risk that publication bias will dominate the estimate obtained from the review and make it unreliable. Even if a study has been published in full, this is no guarantee that its results can be incorporated directly into a meta-analysis without additional information. For example, the reviewer might need to supplement the published data with extra detail on subgroups of participants, further follow-up or the re-inclusion of data from participants mistakenly excluded by the original researcher.

The results to be sought from the original researchers might be aggregate data (for example, by asking them to fill in a table), or data at the level of individual participants. Collection of data from the researchers might make the dataset available for the review more complete, up-to-date and accurate than anything that has been published. It should also facilitate the conduct of standardised analyses across the studies. The collection of individual patient data will provide much greater flexibility for the analyses and, if done in a collaborative way with full participation from the original researchers, such reviews might also benefit from a more rounded interpretation and endorsement of the findings (Stewart & Clarke, 1995).

Statistical analysis

A variety of techniques for combining results from separate studies in meta-analyses are available to the reviewer (Cooper & Rosenthal, 1980; Deeks, 2002). The overriding principle should be that each study is analysed separately and the overall result for the review comes from combining these results from the individual statistics. In this way, participants in one study are only directly compared with others in the same study. By showing the results of the meta-analysis as a forest plot, the relative contribution of each study can be clearly seen, and exploration of differences among the results of studies are made easier (Lewis & Clarke, 2001; Glasziou & Sanders, 2002; Higgins et al., 2003).

In planning and conducting statistical analyses for any review, careful consideration needs to be given to subgroup analyses. One of the rationales for doing a systematic review is to bring together more data than are available for any individual study and it is then tempting to break these data apart again into new subgroups. Caution is needed when doing this because of the possibility that spurious, chance results will be obtained; which will be misinterpreted as being of importance in making decisions about health care (Counsell et al., 1994; Clarke & Halsey, 2001).

Even if there is an a priori reason to expect a subgroup analysis to show something different to the overall result, this is no guarantee that a statistically significant difference is reliable clinically. This is because the more analyses are done, the more likely it is that statistically significant results will be found, even when there is truly no difference between the subgroups. Subgroup analyses in a systematic review should be regarded as a way of showing that the direction of effect is the same across different types of patient or as a generator of a hypothesis for testing in future research. Regardless of whether subgroup analyses are done, it is often more reliable to assume that the overall result is as good, if not a better, estimate of the relative effects of treatments in the particular type of patient than that obtained by looking at the results for just these types of patient in the review. This is because the effect of chance will be smaller for the overall result than it would be on the result in any subgroup.

Systematic reviews might also include sensitivity analyses, which ideally should also be planned in advance. A sensitivity analysis is used to determine how sensitive the results of the systematic review are to the decisions that the reviewer took about how the review was done. They are particularly useful where there is uncertainty about the

choices that a reviewer needs to make. For example, sensitivity analyses could be used to determine the effect of including studies published in languages other than English, of using data from studies assessed to be of poor quality or of choosing one statistical technique over another.

Updating systematic reviews

The intention for Cochrane reviews is that these will be updated at least every 2 years or would be annotated to explain why this has not been done. This desire to keep reviews up to date reflects the fact that they are retrospective research seeking to influence current decisions. Thus, the ideal is that the review includes all relevant research available at the time that it is being used to inform a decision. This is clearly impractical without a process for continually updating reviews as new evidence emerges. Instead, mechanisms for periodic updating are needed, in which new research is sought, appraised and added to the review, if appropriate. The updating process might also serve to maintain the contemporary relevance of the review. This may be especially important if the review uses information that changes over time, such as economic costs, the organisational structures for delivering health care or the processes by which decisions are made about health care.

Appraising and using systematic reviews

Before using a systematic review, those factors that are most important when doing one can be considered in order to assess whether the review is fit for purpose. In some cases, the published review might not contain sufficient information to allow it to be appraised fully but, by bearing these issues in mind, the user of a review should be able to identify whether caution needs to be exercised in its interpretation. One particular reason for the need for caution in interpreting systematic reviews is, as noted above, their retrospective nature. They all rely on factors that are quite often out of the control of the reviewers, since they depend on the research done by other people, in other places and at other times.

The foremost of the potential difficulties is that the review is only possible if the appropriate research has been done. Even if there is a wide consensus that a particular question needs to be addressed in a systematic review, the findings of such a review will be dependent on whether, at some time in the past, other researchers felt likewise and actually did the studies (Alderson & Roberts, 2000). If the studies have been done, then the reviewer would ideally hope to find all of these and to be able to include information and data from them in the review, but this will not always be achievable.

Conclusion

Decisions about health care should be based on the best available evidence. This evidence should be of sufficient quality to be fit for purpose. The evidence needs to be robust against the effects of bias and chance. Systematic reviews, in which as much as possible of the relevant research is sought, appraised, summarised and, if appropriate, meta-analysed, provide the best way to ensure that the necessary evidence is available to people at the time they are making decisions (Tharyan et al., 2005). However, as with all scientific research, whether or not the relevant systematic reviews are available and whether studies are available for these will depend upon the prioritisation of the studies and of the reviews (Chinnock et al., 2005).

References

Alderson, P. & Roberts, I. (2000) Should journals publish systematic reviews that find no evidence to guide practice? Examples from injury research. *British Medical Journal*, 320, 376–7.

Chalmers, I. (1993) The Cochrane Collaboration: preparing, maintaining, and disseminating systematic reviews of the effects of health care. *Annals of the New York Academy of Sciences*, 703, 156–65.

Chinnock, P., Siegfried, N. & Clarke, M. (2005) Is evidence-based medicine relevant to the developing world? Systematic reviews have yet to achieve their potential as a resource for practitioners in developing countries. *PLoS Medicine*, 2 (5), 367–9.

Clarke, M. (2004) Doing new research? Don't forget the old: nobody should do a trial without reviewing what is known. *PLoS Medicine*, 1, 100–2.

Clarke, M. & Halsey, J. (2001) DICE 2: a further investigation of the effects of chance in life, death and subgroup

analyses. *International Journal of Clinical Practice*, 55, 240–2.

Clarke, M.J. & Stewart, L.A. (1994) Systematic Reviews: Obtaining data from randomised controlled trials: how much do we need for reliable and informative meta-analyses? *British Medical Journal*, 309, 1007–10.

Cooper, H.M. & Rosenthal, R. (1980) Statistical versus traditional procedures for summarizing research findings. *Psychological Bulletin*, 87, 442–9.

Counsell, C.E., Clarke, M.J., Slattery, J. & Sandercock, P.A.G. (1994) The miracle of DICE therapy for acute stroke: fact or fictional product of subgroup analysis? *Britsh Medical Journal*, 309, 1677–81.

Deeks, J.J. (2002) Issues in the selection of a summary statistic for meta-analysis of clinical trials with binary outcomes. *Statistics in Medicine*, 21, 1575–600.

Deeks, J.J., Dinnes, J., D'Amico, R. et al. (2003) Evaluating non-randomised intervention studies. *Health Technology Assessment*, 7 (27), 1–183.

Detsky, A.S., Naylor, C.D., O'Rourke, K., McGeer, A.J. & L'Abbe, K.A. (1992) Incorporating variations in the quality of individual randomized trials into metaanalysis. *Journal of Clinical Epidemiology*, 45, 255–65.

Dickersin, K., Min, Y.I. & Meinert, C.L. (1992) Factors influencing publication of research results. Follow-up of applications submitted to two institutional review boards. *Journal of the American Medical Association*, 267, 374–8.

Dickersin, K., Manheimer, E., Wieland, S., Robinson, K.A., Lefebvre, C. & McDonald, S. (2002) Development of the Cochrane Collaboration's CENTRAL Register of controlled clinical trials. *Evaluation and the Health Professions*, 25, 38–64.

Glasziou, P.P. & Sanders, S.L. (2002) Investigating causes of heterogeneity in systematic reviews. *Statistics in Medicine*, 21, 1503–11.

Higgins, J.P.T., Thompson, S.G., Deeks, J.J. & Altman, D.G. (2003) Measuring inconsistency in meta-analyses. *British Medical Journal*, 327, 557–60.

Hopewell, S., Clarke, M., Lusher, A., Lefebvre, C. & Westby, M. (2002) A comparison of handsearching versus MEDLINE searching to identify reports of randomized controlled trials. *Statistics in Medicine*, 21, 1625–34.

Juni, P., Witschi, A., Bloch, R. & Egger, M. (1999) The hazards of scoring the quality of clinical trials for meta-analysis. *Journal of the American Medical Association*, 282, 1054–60.

Lewis, S. & Clarke, M. (2001) Forest plots: trying to see the wood and the trees. *British Medical Journal*, 322, 1479–80.

Pilkington, K., Boshnakova, A., Clarke, M. & Richardson, J. (2005) 'No language restrictions' in database searches: what does this really mean? *Journal of Alternative and Complementary Medicine*, 11, 205–7.

Soares, H.P., Daniels, S., Kumar, A. et al.; Radiation Therapy Oncology Group (2004) Bad reporting does not mean bad methods for randomised trials: observational study of randomised controlled trials performed by the Radiation Therapy Oncology Group. *British Medical Journal*, 328, 22–4.

Stewart, L., Clarke, M., for the Cochrane Collaboration Working Group on meta-analyses using individual patient data (1995) Practical methodology of meta-analyses (overviews) using updated individual patient data. *Statistics in Medicine*, 14, 2057–9.

Tharyan, P., Clarke, M. & Green, S. (2005) How The Cochrane Collaboration is responding to the Asian Tsunami. *PLoS Medicine*, 2 (6), e169.

2 Key Stages and Considerations when Undertaking a Systematic Review: Bladder Training for the Management of Urinary Incontinence

Brenda Roe

Introduction

Systematic reviews are a valuable source of information and help policy makers and clinicians appraise the evidence on which to make decisions. This chapter deals with the systematic identification, appraisal and synthesis of quantitative evidence, notably that from randomised controlled trials (RCTs), and draws on the methods of the Cochrane Collaboration (Green & Higgins, 2005) and others (CRD, 2001; Egger & Davey Smith, 2005), using a systematic review of bladder training for the management of urinary incontinence in adults (Wallace et al., 2004) by way of illustration.

Systematic reviews follow a strict protocol to ensure that as many of the research studies as possible have been considered and original primary studies or trials and papers arising from them are appraised and synthesised in a valid way. The purpose of these systematic methods of review is to minimise bias, provide transparency and enable replication (CRD, 2006). More than one reviewer is involved in independent study inclusion decisions, quality assessment and data extraction, with agreement and consensus reached to avoid individual bias.

Systematic reviews undertaken as part of the Cochrane Collaboration include RCTs, which are recognised as the 'gold standard'. Their reviews adopt an established format and are developed from an initial title and protocol, which are registered with a relevant Cochrane Review Group (CRG). The key stages, procedures and policies are published in each of the CRG websites. Key aspects of Cochrane systematic reviews are that they involve consumers in their production, as well as undergoing scientific and statistical peer review, and are produced according to guidelines in the *Cochrane Handbook for Systematic Reviews of Interventions* (Green & Higgins, 2005). The reviews are published electronically in the Cochrane Library and are disseminated widely via the internet. All reviews are regularly updated.

The Cochrane Incontinence Review Group was established in 1996 and can be accessed via the Cochrane Collaboration website (Grant et al., 2006a). The bladder training review was first published in 1998 (Roe et al., 1998), and two updates have been undertaken (Roe et al., 2000; Wallace et al., 2004). The bladder training review is referred to in this chapter by way of example, but all Cochrane systematic reviews follow the same format and provide an example of robust methods for systematically reviewing quantitative data from RCTs.

Guidelines for undertaking systematic reviews

Textbooks and chapters (Sindhu, 1998; Glasziou et al., 2001; Egger et al., 2005), as well as handbooks

(CRD, 2001; Green & Higgins, 2005), are available as guidance for undertaking a systematic review, and support is provided by Cochrane Review Groups across the globe (see Cochrane Collaboration website for contacts and locations).

Developing a protocol

Developing a protocol is the first step in undertaking a systematic review, as it is with any research endeavour or inquiry. Before a Cochrane systematic review can be undertaken a title needs to be registered and then a protocol developed according to specific criteria; the protocol is then published in the library, having been reviewed by a CRG (see Cochrane Library for examples of protocols). The Cochrane Collaboration runs workshops on 'How to develop a protocol', and these are available for anyone to attend. Irrespective of whether a systematic review is aimed at publication in an academic journal or the Cochrane Library, the protocol formulated needs to include the same considerations and steps (CRD, 2001; Egger & Davey Smith, 2005; Green & Higgins, 2005). According to Egger & Davey Smith (2005), the protocol needs to include seven steps, which relate to:

(1) the research question
(2) inclusion and exclusion criteria
(3) locating studies
(4) selecting studies
(5) assessing the quality of studies
(6) extracting the data
(7) potential analysis and presentation of results.

Steps (4)–(6) all require more than one reviewer to undertake independent assessment and extraction activities and make comparisons to reach agreement and consensus, as required by the systematic review methods of the Cochrane Collaboration to reduce individual bias (Green & Higgins 2005). The systematic review on *Bladder Training for the Management of Urinary Incontinence in Adults* (Wallace et al., 2004) had its protocol first published in 1997 (Roe et al., 1997). This included the background and justification, objectives and hypotheses to be tested, criteria for considering studies (types of studies, participants, interventions, outcome measures), search strategy for identification of studies, inclusion and exclusion criteria for studies, methods for assessment of quality and appropriateness, data extraction, tables of comparisons and analysis.

The objectives and hypotheses tested for the bladder training review on urinary incontinence (whether defined by symptom classification or urodynamic study as indicated by the trialists) are explicit and are measurable (Wallace et al., 2004) (Box 2.1).

Literature searching

The search strategy for identifying relevant studies (trials) should be explicit and included in the methods of the systematic review. Search strategies for identifying controlled trials have developed over recent years, with terms to index RCTs being introduced into the bibliographic databases of MEDLINE and EMBASE. For this purpose, the Cochrane Collaboration examined around 300 000 MEDLINE and EMBASE titles and abstracts, which were then retagged as clinical trials if appropriate. Both databases were examined, as their overlap of journals was around 34% (Smith et al., 1992). The majority of journals in MEDLINE are published in the United States of America, while EMBASE has better coverage of European journals. The retagging of trials in these databases continues, supplemented by manual or hand searches of journals, conference proceedings, other sources and specialised databases. The results of retagging and hand searches have been included in The Cochrane Controlled Trials Register in the Cochrane Library, which includes over 250 000 trials and is the best single source of published studies. Searches of MEDLINE and EMBASE are still recommended, along with other specialised databases, conference abstracts, monographs and references in review articles. Hand searching is also recommended as part of the search strategy, as is identifying unpublished studies by contacting lead investigators in order to remove publication bias.

Each CRG has explicit search strategies and those for incontinence are available on the Cochrane Incontinence Review Group website (Grant et al., 2006a), and include electronic searches of the Cochrane Central Register of controlled trials (CENTRAL), MEDLINE and the Cumulative Index of Nursing and Allied Health Literature (CINAHL), and hand searching of journals and conference

Box 2.1 Objectives and hypotheses tested by the Cochrane review on bladder training for urinary incontinence in adults (reproduced from Wallace et al. (2004) with permission from J. Wiley & Sons, Chichester and Sheila Wallace).

Objectives

To assess the effects of bladder training on urinary incontinence, however the diagnosis is made.
The following hypotheses were tested:

(1) Bladder training is better than no bladder training for the management of urinary incontinence.
(2) Bladder training is better than other treatments (such as conservative or pharmacological) for the management of urinary incontinence. This hypothesis will be tested by looking at the following comparisons:
    (a) bladder training compared with anticholinergic drugs;
    (b) bladder training compared with adrenergic agonist drugs;
    (c) bladder training compared with other drugs (non-anticholinergic, non-adrenergic agonist drugs);
    (d) bladder training compared with other behavioural /physical/psychological treatments;
    (e) bladder training compared with surgical management;
    (f) bladder training compared with medical devices;
    (g) bladder training compared with other intervention.
(3) Combining bladder training with another treatment (such as conservative or pharmacological) is better than the other treatment alone. This hypothesis will be tested by looking at the following comparisons:
    (a) bladder training combined with pharmacological treatment compared with that of pharmacological treatment alone;
    (b) bladder training combined with a non-pharmacological treatment compared with that of non-pharmacological treatment alone.

proceedings to identify published and unpublished trials. There is also an attempt to hand search journals that are not in English. Topic specific search terms for urinary incontinence were combined (with the Boolean operator AND) with the randomised controlled trials methodology terms. The revised CRG design methodology search strategy for randomised controlled trials in PubMed is shown in Box 2.2, while search terms specific to bladder training are given in Box 2.3. It is advisable when developing a search strategy to consider electronic and hand searching, and the Medical Subject Heading (MeSH) terms and keywords to be used.

Initial development, testing and refinement are important unless 'standardised' searching according to Cochrane and a CRG are used. It is advisable to finalise the search strategy with a librarian or information scientist. The search strategy for identification of studies is a standard entry in all Cochrane reviews. For the bladder training review extra specific searches also included reference lists of relevant articles and contact with investigators for information on other possible trials that were published or unpublished, and no year or language limits were set (Wallace et al., 2004).

Publication bias

Historically, there was a tendency for only trials that found statistically significant findings to be published (Sindhu, 1998). There is also evidence that authors from high-prestige organisations are more likely to have their studies published than those from lower-prestige organisations (Peters & Ceci, 1982; Egger et al., 2005). These publication policies can influence what studies are published and represent publication bias. Publication bias can threaten the validity of the meta-analysis within systematic reviews as not all results or findings are available or known about and can be compared, and this can distort the results. In order to reduce publication bias, every effort needs to be made to locate all trials on a particular subject when undertaking a systematic review. This can be addressed by hand searching or electronic searching of conference

Box 2.2 Cochrane highly sensitive search strategy for identifying reports of randomised controlled trials in PubMed (2005 revision) (source: Glanville et al., 2006, permission for reproduction granted by Carol Lefebvre of the Cochrane Collaboration and colleagues and the Medical Library Association).

1. clinical trial [pt] need to explode in OVID as automatic in PubMed
2. randomized [tiab]*
3. placebo [tiab]
4. dt [sh]*
5. randomly [tiab]
6. trial [tiab]
7. groups [tiab]
8. #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7
9. animals [mh]
10. humans [mh]
11. #9 NOT (#9 AND #10)
12. #8 NOT #11

*[pt] denotes Publication Type; [mh] denotes Medical Subject Heading (MeSH); [tiab] denotes a word in the title or abstract; [sh] denotes a subheading.

Sets 9 to 11 of the strategy capture animal studies that are also not human studies, and allow these records to be safely excluded from the search, while returning records that are not indexed as either human or animal studies, as these may be relevant.

Box 2.3 Cochrane Incontinence Review Group keyword system search terms for bladder training for use in the Cochrane Central Register for Controlled Trials (CENTRAL) (reproduced with permission from Sheila Wallace, Search Coordinator for the Cochrane Incontinence Review Group).

Search terms for CENTRAL

Bladder NEAR/2 (train or retrain* or educat* or re-educat* or drill) in All Fields in The Cochrane Central Register of Controlled Trials (CENTRAL)

Key: * = truncation symbol; NEAR/2 = search for terms within two words of bladder.

proceedings, and contacting directly principal investigators and organisations known to fund work in the related area or that have an interest in the particular subject. Another way to address this is for all trials to be entered on a register when they commence. Registers have been set up and can be accessed via the Cochrane Collaboration's Register of Registers. Others can be accessed via the internet, such as the metaRegister of Controlled Trials published by *Current Science* (Lefebvre & Clarke, 2005).

Inclusion criteria

Explicit inclusion criteria for studies that are to be included in a systematic review are required at the start. However, decisions to include or exclude studies are to a certain extent subjective, despite having explicit criteria. Methods for undertaking systematic reviews recommend using two observers to check the eligibility of studies for inclusion, with discussion and consensus as to those that should

be included. If agreement is not possible, then a third reviewer can also be involved (Egger & Davey Smith, 2005). Systematic reviews should involve more than one reviewer, which is a requirement for Cochrane reviews, and ideally reviewers should have a variety of multidisciplinary backgrounds and international perspectives.

In the bladder training review, inclusion criteria were pre-specified for types of studies, participants and interventions (Wallace et al., 2004). Studies included all randomised or quasi-randomised controlled trials that included bladder training for the treatment of urinary incontinence. Urinary incontinence was defined and diagnosed by the trialists either by symptom classification or by urodynamics. Subjects were all adult men and women with urinary incontinence, and the term adult was accepted as defined by the trialists. Studies that were eligible also had to include at least one trial group receiving bladder training, even if explicit descriptions of bladder training were not described. As long as the term 'bladder training' was stated, studies that fulfilled the above criteria were eligible. Bladder drill, bladder re-training and bladder re-education were accepted as being synonymous with bladder training. Studies were included if the following specific terms were not used but they comprised the intervention:

- Mandatory schedule or a self-schedule which increased the time interval between voids, as a minimum, and
- Participant education, and
- Positive reinforcement and follow-up

If the intervention was unclear, then trialists were contacted for clarification. No restrictions were set for where bladder training took place and this could include out-patient, in-patient or home settings, although these locations were not compared as to their effectiveness and nor was bladder training being undertaken by different healthcare professionals (Wallace et al., 2004).

Exclusion criteria

Studies excluded from a systematic review and the reasons why are also explicit and are published within individual reviews in the Cochrane Library. Exclusion criteria are also stated in advance in the protocol and are adhered to. For instance, in relation to the bladder training review, studies that did not fulfil the above inclusion criteria were excluded. Those that also described bladder training as it related to the clamping or removal of urinary catheters were excluded. If trials did not include mention of a mandatory or self-schedule, they too were excluded. If an additional intervention was added to supplement bladder training, such as pelvic floor muscle exercise training (PFMT) compared to no treatment, 'usual care' or bladder training alone, these trials were also excluded as it is not possible to assess the direct effects of bladder training (Wallace et al., 2004). This illustrates the importance of having exclusion criteria as well as inclusion criteria for studies when undertaking a rigorous systematic review, and having a minimum of two reviewers or observers to assess which studies are included and excluded by consensus.

Quality assessment

Once studies have been selected for inclusion, an assessment is made about the quality of their design in relation to randomisation and blinding of subjects, people undertaking the intervention and those measuring outcomes. More than one reviewer should undertake assessment of the quality of each included study independently, with agreement reached by consensus. Although randomised controlled trials are the gold standard and provide the best evidence for efficacy of interventions, they are still vulnerable to bias. The quality of a trial can influence the effect size (Egger & Davey Smith, 2005). Inadequate concealment or blinding of randomisation and group allocation can lead to larger treatment effects. Treatment effects can be overestimated when 'intention to treat' analyses are not undertaken and subjects withdrawing or not adhering to the intervention are not included in the analysis. However, there is a divergent view within the Cochrane Collaboration that intention to treat analysis should only include those who received and completed the treatment and exclude those that withdrew (S. Wallace, personal communication). Based on the bladder training review, it is apparent that the older trials did not include sufficient detail of how random allocation was undertaken and whether concealment was achieved.