Bridging the gap between research and clinical practice in modern pediatrics

Boluyt, N.

Publication date
2007
Document Version
Final published version

BRIDGING THE GAP BETWEEN RESEARCH AND CLINICAL PRACTICE IN MODERN PEDIATRICS Nicole Boluyt

BRIDGING THE GAP BETWEEN RESEARCH AND CLINICAL PRACTICE IN MODERN PEDIATRICS

Primary Studies, Systematic Reviews and Clinical Practice Guidelines

Nicole Boluyt

Bridging the gap between research and clinical practice in modern pediatrics

Primary Studies, Systematic Reviews and
Clinical Practice Guidelines

Bridging the gap between research and clinical practice. Primary Studies, Systematic Reviews and Clinical Practice Guidelines. Thesis, University of Amsterdam, The Netherlands

Lay-out: Annelies Bast
Cover: Annelies Bast
Printed by: Tailormade, Buren

ISBN: 978-90-9021719-2

Bridging the gap between research and clinical practice in modern pediatrics

Primary Studies, Systematic Reviews and Clinical Practice Guidelines

Academisch proefschrift

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam op gezag van
de Rector Magnificus prof. mr. P.F. van der Heijden
ten overstaan van een door het college
voor promoties ingestelde commissie, in het openbaar
te verdedigen in de Aula der Universiteit
op dinsdag 24 april 2007, te 12.00 uur door

Nicole Boluyt

geboren te Wassenaar

Promotiecommissie

Promotor
Prof. dr. M. Offringa

Overige leden
Prof. dr. W.M.C. van Aalderen
Prof. dr. P.M.M. Bossuyt
Dr. P.L.P. Brand
Prof. dr. H.N. Caron
Prof. dr. R.P.T.M. Grol
Prof. dr. J.H. Kok
Dr. R.J.P.M. Scholten

Faculteit der Geneeskunde

Voor mijn ouders

Contents

Chapter 1	Introduction and outline of the thesis	9
Chapter 2	The usefulness of systematic review search strategies in finding child health systematic reviews in MEDLINE *Submitted*	23
Chapter 3	Neurodevelopment after neonatal hypoglycemia: a systematic review and design of an optimal future study *Pediatrics. 2006 Jun;117(6):2231-43*	39
Chapter 4	State of the evidence on acute asthma management in children: a critical appraisal of systematic reviews *Submitted*	67
Chapter 5	Fluid resuscitation in neonatal and pediatric hypovolemic shock: a Dutch Pediatric Society evidence-based clinical practice guideline *Intensive Care Med. 2006 Jul;32(7):995-1003*	89
Chapter 6	Implementation and evaluation of a pediatric evidence-based guideline for fluid resuscitation in hypovolemia *Submitted*	107
Chapter 7	Quality of evidence-based pediatric guidelines *Pediatrics. 2005 May;115(5):1378-91*	125
Chapter 8	General discussion and recommendations	155
	Summary	185
	Samenvatting voor de leek	191
	Co-authors	201
	Dankwoord	203
	Curriculum vitae	207

CHAPTER 1

Introduction and outline of the thesis

**''The fool doth think he is wise,
but the wise man knows himself to be fool"
Shakespeare**

Introduction

Doctors' decisions may have profound effects on their patients' health. Therefore these decisions should be of the best possible quality. Clinicians often have to make rapid decisions, either because of medical emergency or because they see so many patients in a limited time.

Psychologists have shown that rapid decision making is aided by heuristics, i.e. strategies that provide shortcuts to quick decisions, but they also remark that these heuristics are frequently misleading. Good decision making in clinical practice is impeded by the fact that we often fall prey to various cognitive biases. One of them is overconfidence. Research has shown that almost all of us are more confident about our judgments than we should be.[1] The dangers are obvious: doctors who overestimate their abilities in the management of a condition may continue to prescribe suboptimal treatment, unaware that their management could be improved. Therefore, it is of critical importance to be aware of the limits of one's knowledge and to assure that this knowledge is kept up to date. In child health care, this is painfully shown by the following example:

> Sudden unexpected unexplained infant death, now known as sudden infant death syndrome (SIDS), was recognized as a major cause of infant death throughout the 20th century. In the early sixties and seventies, Dr Benjamin Spock, a pediatrician from the United Sates, wrote the most widely recommended handbook for parents ever written and over 19 million copies were sold.[2] Parents were advised to place their baby on the front, as putting a baby on the back was considered to enhance the risk of suffocation in case the baby vomited. In 1988 the first overview of studies on the effect of sleeping positions showed that infants put on the front had an increased risk of SIDS. A subsequent systematic historical review showed that in 1970 there was already sufficient evidence that putting infants to sleep on the front was likely to be harmful.[3] Earlier recognition of the evidence might have prevented over 50.000 infant deaths in Europe, the USA and Australasia. After a campaign was started in the early 1990s to put infants on their back, the incidence of SIDS decreased from 2,5 per 1000 newborns in 1990 to 0,5 per 1000 newborns in 1992.

Keeping up to date in medicine in the 21st century is exceptionally challenging. More than 2 million new scientific papers are published each year. Keeping up with general pediatric literature alone would require reading at least 5 journal articles each day, 365

days a year.[4] Keeping up to date on medical developments is an explicit component of professionalism and mandated in the form of continuing medical education requirements. Unfortunately, educational updates with traditional continuing medical education have been proven insufficient to significantly affect either physicians' practices or the health outcomes of their patients.[5] Although many providers believe that they do apply the latest evidence in the care of their patients, there are ample data to suggest that in fact they do not. Considerable lags exist between the confirmation that given therapies or tests are helpful and their translation into clinical practice.[6] Moreover, ineffective therapies continue to be used despite convincing evidence that they do more harm than good.[7]

Evidence Based Pediatrics and Child Health

The term evidence-based medicine (EBM) was coined in 1991 by Sackett and colleagues at McMaster University in Hamilton, Ontario, and has since then rapidly evolved.[8] Evidence-based pediatrics (EBP) is defined as 'the integration of clinical information obtained from a patient with the best evidence available from clinical research and experience, and the application of this knowledge to the prevention, diagnosis and management of disease in the child'.[4]

But surely, physicians have always used evidence in caring for their patients, so what is new? What is new is our recognition that, despite our best intentions, dramatic variations in care occur from physician to physician, even in caring for similar patients. This implies that we have not been using the current best evidence in the care of our patients, even though we intended to do so. The most important reason for practicing EBP is to improve the quality of care for children through the identification and promotion of effective treatments, and the elimination of those that are ineffective or harmful.

The practice of EBP involves several steps. It begins and ends with the individual patient. Once the physician has obtained clinical information, the next step is to identify the important clinical questions arising from the individual child and structure these using a framework (the so called P.I.C.O. format) which will make them likely to be answerable by the literature. The search for available evidence to address these questions is followed by critical appraisal to determine the validity of the evidence, i.e. the likelihood that results are not flawed, and its relevance or usefulness (applicability) to the patient in question. Application of evidence to decision making in clinical practice must also take into consideration patient preference, family circumstance and issues of quality of life. EBP does not discount the need for clinical skills. Rather, it recognizes that both knowledge of the evidence and clinical expertise are necessary and that neither alone is sufficient for the best practice of medicine. Evaluation of the process, both with regard to patient outcome and educational benefit to the clinician, should be the final step in the practice of EBP.[4]

Getting good research evidence into pediatric practice takes more than following all steps of evidence-based medicine, which consists of applying the results of a single research report to an individual patient. It has been recognized for some time that clinicians need summaries of all the available evidence on a specific topic, presented in systematic literature reviews and evidence-based practice guidelines.

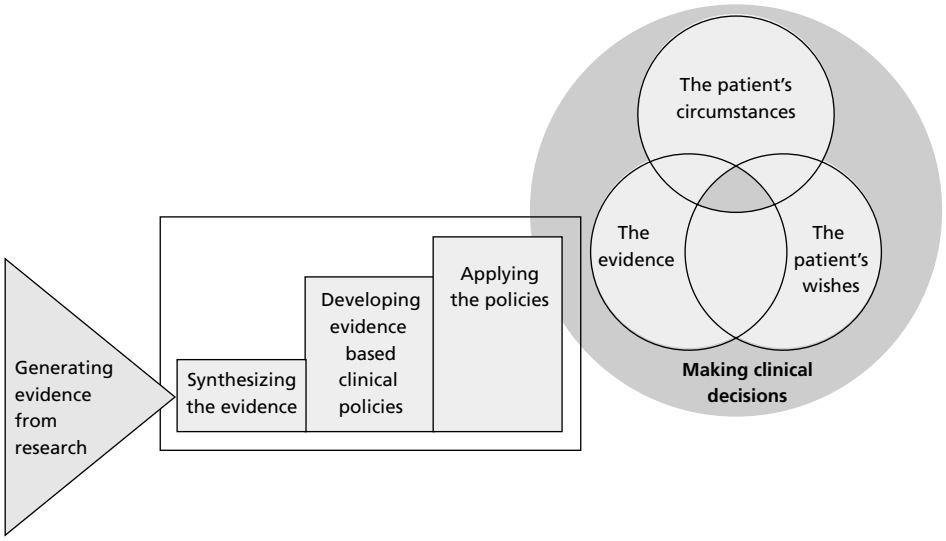
Getting evidence into practice

For many health conditions, a gap exists between what medical science has shown to be effective practice and what is actually done.[3;9] This is clearly shown by the following example.

> Dexamethasone given to mothers in premature labor is known to prevent neonatal respiratory distress syndrome by accelerating fetal lung maturation. Its efficacy was already demonstrated in clinical trials published in the early seventies. Despite a significant beneficial effect being confirmed in a number of further trials and a meta-analysis published in 1990, the uptake of this technology was astonishingly low.[9;10] It was estimated in 1995 that only 12-18% of eligible mothers received this treatment in the United States.

It is now recognized that several steps are needed to harness research evidence for healthcare practice (Figure 1).[11] These steps include synthesizing the evidence into systematic reviews; developing clinical policy from the evidence into evidence-based practice guidelines; and applying the policy, i.e. implementation of guidelines.

Figure 1 The path from the generation of evidence to the application of evidence[11]



Systematic reviews

As pointed out earlier, biomedical scientific knowledge is expanding far too quickly for any individual clinician or researcher to keep up. Clinicians need summaries of all best evidence on a given clinical topic. Traditionally this was done by providing *narrative* reviews of topic areas by experts in the field. However, this is more and more abandoned since the methods of such a review are not transparent. Authors may cite studies selectively which support their own opinion and fail to cite other studies which provide evidence which conflicts with their opinion. *Systematic* reviews of the literature address a well-defined clinical question, use an explicit search strategy to locate *all* relevant evidence, evaluate the retrieved studies using prospectively defined methodological criteria, and synthesize the results according to a pre-specified method.[12;13]

The Cochrane Collaboration is an international network of individuals, mostly clinicians, whose goal is to undertake systematic reviews addressing important therapeutic questions, using standardized methodology for searching and appraising the literature and reporting the results. The Cochrane Library is the prime source for systematic reviews, but it does not contain all systematic reviews published in peer reviewed journals. Large bibliographic databases such as MEDLINE can be used to identify other systematic reviews.

Well conducted systematic reviews are increasingly seen as providing the best evidence to guide clinical practice and should be the cornerstone for the recommendations of evidence-based practice guidelines. In addition, they guide future research into the information gaps that are identified through the review process. Several sources of bias have been identified that may limit the validity of systematic reviews and therefore checklists have been provided to enable readers to assess the quality of a review article.[14] Several studies have shown that most systematic reviews and meta-analyses on health care interventions in adult patients that were published in peer reviewed journals have methodological deficiencies that limit their validity and the applicability of their results in practice.[15-20] So far, there is no information regarding the validity and applicability of systematic reviews in child health care.

Evidence-based clinical practice guidelines

The definition of a clinical guideline formulated by the Dutch Institute for Healthcare Improvement CBO is 'a guideline is a document with recommendations, guidance and instructions to support daily practice in health care, based on the results of scientific research and the consequent discussion and formation of opinion, aimed at the explicit statement of good medical practice'.[21]

The aim of clinical guidelines is to improve quality of care by translating new research findings into clinical practice.[22] Well conducted guidelines have been shown to have the power to improve patient outcomes. However, several other studies have also shown that the quality of (adult) evidence-based guidelines is generally poor.[23-27]

Initially, guidelines developed by pediatricians were based on the experience and opinions of a number of experts. In the 1990s, scientific requirements for clinical research and guidelines became stricter, partly due to the evidence-based medicine movement. The Dutch Board for Pediatricians (NVK) started an evidence-based guideline development programme in 2000 and since then 5 pediatric guidelines have been developed; neonatal resuscitation, volume resuscitation, pain management, SIDS and mental retardation.

Guideline implementation

Passive dissemination of clinical practice guidelines does not result in behavior change.[12] Specific strategies to implement research-based recommendations appear necessary to ensure practice change. Several systematic reviews including hundreds of studies have taught us that there is no 'magic bullet' for implementation success.[28-31] Based on systematic reviews published by the Cochrane Effective Practice and Organization of Care Group (EPOC) implementation strategies can be categorized into three groups showing consistent, variable, or little or no effectiveness. Those interventions that *consistently have shown effectiveness* include interactive educational meetings, educational outreach visits, reminders, and multifaceted interventions (defined as a minimum of two combined interventions). Interventions that have shown *variable effectiveness* include audit and feedback, the use of local opinion leaders, local consensus processes, and patient-mediated interventions. Interventions that have consistently showing *little or no effect* are didactic educational meetings (lecture-format) and educational materials.[30-33] So far, there is very little experience with the implementation of pediatric evidence-based guidelines.

Practicing evidence-based pediatrics

Although evidence-based medicine was first launched in the early nineties, the practice of evidence-based pediatrics and child health is relatively new. It is not yet widely implemented in the Netherlands or abroad. The main reason for this slow uptake is that practitioners find it challenging to find and appraise the evidence they need for daily clinical decision making. In fact, they are struggling with the various aspects of the utility of evidence. The utility of evidence for a given clinical problem can be summarized by the following equation:

$$\text{Utility of evidence} = \frac{VxQ}{W},$$

in which V is the volume of the evidence, Q is the quality of the evidence, and W is the amount of work done to find evidence.[34]

Actually, the utility of evidence for most child health topics is not known. To evaluate the utility of evidence in pediatrics, the following questions need to be addressed:

1 How and where do we find the evidence efficiently? (related to W)
2 What is the volume of the evidence? (related to V)
3 What is the quality of the evidence? (related to Q)

This thesis

This thesis is concerned with challenges that occur while practicing EBP – the efficient and correct application of research evidence in pediatric clinical practice. These include the above-mentioned EBM components: searching for evidence, establishing the volume of the evidence, and determining the quality of the evidence. It focuses on the three steps between an original study as published in the literature and clinical decision making (Figure 1). These parts are systematic reviews, clinical practice guidelines and implementation of guidelines (Table 1).

The clinical topics presented in this thesis are 'question driven': they are topic areas of specific importance identified by practicing pediatricians. Below these problems will be described briefly. The subsequent chapters of this thesis will address these questions in detail.

Outline of this thesis (see Table 1)

There is currently a considerable need for compiled, synthesized evidence in child health. The Cochrane Child Health Field (www.cochranechildhealth.ualberta.ca), a partnership of all child health care providers within the Cochrane Collaboration, has developed a list of *priority topics* in pediatrics, identified by users in the field, which need systematic reviews. Yet, at this moment it is not clear whether systematic reviews on these topics already exist – as it is not clear how to locate systematic reviews in children efficiently.

Problem 1
How to find existing child health systematic reviews in Medline?

Chapter 2 describes a study in which the usefulness of existing systematic review search strategies, combined with a child filter, to find child health systematic reviews in MEDLINE is evaluated by applying them to a reference standard of child health systematic reviews. Data on sensitivity, number of records retrieved and precision are calculated in the whole of PubMed and by combining filters with seven child health priority topics. Recommendations for clinicians are given on where and how to search for child health systematic reviews.

Hypoglycemia is the most common metabolic problem in neonatal medicine.[35] Still, there is much controversy about the definition of a 'safe' blood glucose concentration,

i.e. a value above which there is no risk of long-term neurodevelopmental impairment, and as a result clinical practice varies widely.[36;37]

Problem 2
What is the neurodevelopmental outcome after neonatal hypoglycemia?

Chapter 3 describes a prognostic systematic review of observational studies that examine the neurodevelopmental outcome after neonatal hypoglycemia in the first week of life. Building on the strengths and weaknesses of existing studies a proposal for an 'optimal future study' is developed.

Acute asthma is a common reason for coming to emergency departments.[38] In this field, systematic reviews have already gained popularity as a way of coping with increasing amounts of information about new devices and drugs as they synthesize large amounts of research evidence to help bridge the gap between evidence from single studies and clinical practice. Yet, if clinicians are to have confidence that the results of systematic reviews can be used to guide clinical practice and the research agenda, then these reviews need to be of high quality.

Problem 3
What is the quality of pediatric systematic reviews on acute asthma management in order to guide clinical practice and the research agenda?

Chapter 4 evaluates clinical, methodological and reporting aspects of systematic reviews on the treatment of acute asthma in children. In addition, it is assessed how data on children are reported in 'mixed population' reviews, and Cochrane reviews are compared with reviews published in peer reviewed journals. Recommendations for future trials and systematic reviews are presented.

Considerable controversy and variation in clinical practice exists on whether colloids or crystalloids should be used for fluid resuscitation in pediatric hypovolemic shock.[39] This uncertainty resulted in a plea for a Dutch evidence-based guideline.

Problem 4
What is the best first choice fluid for resuscitation of hypovolemic neonates and children?

Chapter 5 describes the development of the first national pediatric evidence-based guideline in the Netherlands on fluid resuscitation in pediatric hypovolemic shock.

Practice guidelines are not used in daily practice unless actively implemented pursuing consolidation of behavior change.[30;33] Implementing guidelines is a great challenge: evidence about the most effective and efficient guideline implementation strategies in various different circumstances is still lacking. Given a new practice guideline that was developed with funding of our national Pediatric Society, we identified

Problem 5
How to implement a pediatric clinical practice guideline?

Chapter 6 describes the methodology used for implementation of a national evidence-based guideline and the success of its implementation according to different indicators. Based on what we learnt while implementing this specific guideline, general advise for future guideline implementation projects is presented.

Pediatricians in the Netherlands are getting more and more interested in using evidence-based practice guidelines. They were asked to prioritize the five most urgent topics for future guideline development. Yet, as the process of developing pediatric evidence-based guidelines is time-consuming and needs extensive resources and as the number of existing international pediatric evidence-based guidelines is rapidly increasing, we might adjust these existing guidelines for local use.

Problem 6
What is the quality of existing international pediatric evidence-based guidelines and how can they be adjusted for local use?

Chapter 7 describes the volume and quality of international evidence-based guidelines. Three pediatrician-reviewers appraised the available guidelines on ten priority topics identified by Dutch pediatricians using the Appraisal of Guidelines for Research and Evaluation (AGREE) instrument.[40] Furthermore, we give advise on how to adjust 'high quality' guidelines for local use.

Chapter 8 discusses several challenges in practicing Evidence-Based Pediatrics and Child Health. The discussion of these challenges builds upon what we have learned in conducting the studies described in this thesis, integrated with knowledge that has been developed in this field during the last few years. These challenges are:

1 To use all existing evidence to guide clinical practice and the research agenda
2 To avoid unnecessary duplication of effort
3 To improve the quality of evidence in pediatrics
4 To effectively translate research findings into pediatric clinical practice
5 To enhance the understanding of research methods among its end-users

Finally, all the lessons we have learnt are translated into recommendations for future a) primary studies, b) systematic reviews, and c) clinical practice guidelines in 21st century pediatrics.

Table 1 Thesis outline

Question	Clinical topic	Object of study	Design	Chapter
A SYSTEMATIC REVIEWS				
How to find child health systematic reviews in MEDLINE?	Asthma Obesity Constipation UTI* Enuresis Bronchiolitis ADHD±	Existing systematic review search strategies combined with a child filter	Accuracy study	2
What is the neurodevelopmental outcome after neonatal hypoglycemia?	Neonatal hypoglycemia	Cohort studies measuring neuro-developmental outcome	Systematic review	3
What is the current state of the evidence on acute asthma management?	Acute asthma	Systematic reviews of acute asthma interventions	Systematic review of reviews	4
B GUIDELINES				
What is the first choice fluid for resuscitation of hypovolemic neonates and children?	Fluid resuscitation in critically ill children	RCTs comparing colloids and crystalloids and 'other considerations' in the decisions which fluid to use	Guideline development	5
How to implement an evidence-based practice guideline?	Fluid resuscitation in critically ill children	Clinical practice change	Guideline implementation	6
What is the quality of existing pediatric evidence-based guidelines?	Constipation UTI* Head Injury Diabetic ketoacidosis Sedation for procedures Vesicoureteral reflux Antiemetics for patients receiving chemotherapy Fever of unknown origin	Evidence-based guidelines	Guideline Quality and Utility Assessment	7

* Urinary Tract Infection

± Attention Deficit Hyperactivity Disorder

Reference List

1 Klein JG. Five pitfalls in decisions about diagnosis and prescribing. BMJ 2005; 330 (7494):781-783.

2 Spock B. Baby and Child Care. London: The Bodley Head; 1958.

3 Gilbert R, Salanti G, Harden M, See S. Infant sleeping position and the sudden infant death syndrome: systematic review of observational studies and historical review of recommendations from 1940 to 2002. Int J Epidemiol 2005; 34(4):874-887.

4 Moyer VA. Evidence Based Pediatrics and Child Health. Second Edition ed. BMJ Books; 2004.

5 Davis DA, Thomson MA, Oxman AD, Haynes RB. Changing physician performance. A systematic review of the effect of continuing medical education strategies. JAMA 1995; 274(9):700-705.

6 Antman EM, Lau J, Kupelnick B, Mosteller F, Chalmers TC. A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts. Treatments for myocardial infarction. JAMA 1992; 268(2):240-248.

7 Nyquist AC, Gonzales R, Steiner JF, Sande MA. Antibiotic prescribing for children with colds, upper respiratory tract infections, and bronchitis. JAMA 1998; 279(11):875-877.

8 Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. BMJ 1996; 312(7023):71-72.

9 Gilbert R, Logan S. Future prospects for evidence-based child health. Arch Dis Child 1996; 75(6):465-468.

10 Crowley P. Prophylactic corticosteroids for preterm birth. Cochrane Database Syst Rev 2000;(2):CD000065.

11 Haynes B, Haines A. Barriers and bridges to evidence based clinical practice. BMJ 1998; 317(7153):273-276.

12 Bero LA, Grilli R, Grimshaw JM, Harvey E, Oxman AD, Thomson MA. Closing the gap between research and practice: an overview of systematic reviews of interventions to promote the implementation of research findings. The Cochrane Effective Practice and Organization of Care Review Group. BMJ 1998; 317(7156):465-468.

13 Oxman AD, Cook DJ, Guyatt GH. Users' guides to the medical literature. VI. How to use an overview. Evidence-Based Medicine Working Group. JAMA 1994; 272(17):1367-1371.

14 Oxman AD. Checklists for review articles. BMJ 1994; 309(6955):648-651.

15 Assendelft WJ, Koes BW, Knipschild PG, Bouter LM. The relationship between methodological quality and conclusions in reviews of spinal manipulation. JAMA 1995; 274 (24):1942-1948.

16 Choi PT, Halpern SH, Malik N, Jadad AR, Tramer MR, Walder B. Examining the evidence in anesthesia literature: a critical appraisal of systematic reviews. Anesth Analg 2001; 92(3):700-709.

17 Delaney A, Bagshaw SM, Ferland A, Manns B, Laupland KB, Doig CJ. A systematic evaluation of the quality of meta-analyses in the critical care literature. Crit Care 2005; 9(5):R575-R582.

18 Dixon E, Hameed M, Sutherland F, Cook DJ, Doig C. Evaluating meta-analyses in the general surgical literature: a critical appraisal. Ann Surg 2005; 241(3):450-459.

19 Jadad AR, McQuay HJ. Meta-analyses to evaluate analgesic interventions: a systematic qualitative review of their methodology. J Clin Epidemiol 1996; 49(2):235-243.

20 Jadad AR, Moher M, Browman GP, Booker L, Sigouin C, Fuentes M et al. Systematic reviews and meta-analyses on treatment of asthma: critical evaluation. BMJ 2000; 320(7234):537-540.

21 http://www.cbo.nl/product/richtlijnen/

22 Grol R. Improving the quality of medical care: building bridges among professional pride, payer profit, and patient satisfaction. JAMA 2001; 286(20):2578-2585.

23 Cluzeau FA, Littlejohns P, Grimshaw JM, Feder G, Moran SE. Development and application of a generic methodology to assess the quality of clinical guidelines. Int J Qual Health Care 1999; 11(1):21-28.

24 Graham ID, Beardall S, Carter AO, Glennie J, Hebert PC, Tetroe JM et al. What is the quality of drug therapy clinical practice guidelines in Canada? CMAJ 2001; 165(2):157-163.

25 Grilli R, Magrini N, Penna A, Mura G, Liberati A. Practice guidelines developed by specialty societies: the need for a critical appraisal. Lancet 2000; 355(9198):103-106.

26 Shaneyfelt TM, Mayo-Smith MF, Rothwangl J. Are guidelines following guidelines? The methodological quality of clinical practice guidelines in the peer-reviewed medical literature. JAMA 1999; 281(20):1900-1905.

27 Ward JE, Grieco V. Why we need guidelines for guidelines: a study of the quality of clinical practice guidelines in Australia. Med J Aust 1996; 165(10):574-576.

28 Grimshaw J, Eccles M, Thomas R, MacLennan G, Ramsay C, Fraser C et al. Toward evidence-based quality improvement. Evidence (and its limitations) of the effectiveness of guideline dissemination and implementation strategies 1966-1998. J Gen Intern Med 2006; 21 Suppl 2:S14-20.:S14-S20.

29 Grimshaw JM, Shirran L, Thomas R, Mowatt G, Fraser C, Bero L et al. Changing provider behavior: an overview of systematic reviews of interventions. Med Care 2001; 39(8 Suppl 2):II2-45.

30 Grimshaw JM, Thomas RE, MacLennan G, Fraser C, Ramsay CR, Vale L et al. Effectiveness and efficiency of guideline dissemination and implementation strategies. Health Technol Assess 2004; 8(6):iii-72.

31 Oxman AD, Thomson MA, Davis DA, Haynes RB. No magic bullets: a systematic review of 102 trials of interventions to improve professional practice. CMAJ 1995; 153(10): 1423-1431.

32 Bero LA, Grilli R, Grimshaw JM, Harvey E, Oxman AD, Thomson MA. Closing the gap between research and practice: an overview of systematic reviews of interventions to

promote the implementation of research findings. The Cochrane Effective Practice and Organization of Care Review Group. BMJ 1998; 317(7156):465-468.

33 Grimshaw JM, Shirran L, Thomas R, Mowatt G, Fraser C, Bero L et al. Changing provider behavior: an overview of systematic reviews of interventions. Med Care 2001; 39(8 Suppl 2):II2-45.

34 Martin M.W., Sell J. The Marginal Utility of Information: Its Effects upon Decision-Making. The Sociological Quarterly 1980; 21(2):233.

35 Cornblath M. Neonatal hypoglycemia 30 years later: does it injure the brain? Historical summary and present challenges. Acta Paediatr Jpn 1997; 39 Suppl 1:S7-11.

36 Cornblath M, Schwartz R, Aynsley-Green A, Lloyd JK. Hypoglycemia in infancy: the need for a rational definition. A Ciba Foundation discussion meeting. Pediatrics 1990; 85(5):834-837.

37 Cornblath M, Hawdon JM, Williams AF, Aynsley-Green A, Ward-Platt MP, Schwartz R et al. Controversies regarding definition of neonatal hypoglycemia: suggested operational thresholds. Pediatrics 2000; 105(5):1141-1145.

38 Mannino DM, Homa DM, Pertowski CA, Ashizawa A, Nixon LL, Johnson CA. Surveillance for asthma - United States, 1960-1995. Morbidity and Mortality Weekly Report. CDC Surveillance Summaries 1998; 47:1-27.

39 Offringa M. Excess mortality after human albumin administration in critically ill patients. Clinical and pathophysiological evidence suggests albumin is harmful. BMJ 1998; 317(7153):223-224.

40 Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: the AGREE project. Qual Saf Health Care 2003; 12(1):18-23.

CHAPTER 2

The usefulness of systematic review search strategies in finding child health systematic reviews in MEDLINE

Nicole Boluyt, Lisa Tjosvold, Carol Lefebvre, Terry P. Klassen, Martin Offringa

Submitted for publication

ABSTRACT

Objective To determine the sensitivity and precision of existing search strategies for retrieving child health systematic reviews in MEDLINE using PubMed.

Methods We identified existing search strategies for systematic reviews, combined them with a filter that identifies articles relevant to child health, and applied them in MEDLINE to a reference set of child health systematic reviews. Outcome measures were total number of records retrieved, sensitivity, and precision. Sensitivity was defined as the proportion of systematic reviews retrieved from our reference standard. Precision was defined as the proportion of 'true' systematic reviews retrieved from the filters' search yield from MEDLINE and from searches combining these filters with seven child health priority topics.

Results We tested 9 search filters. Sensitivity of the systematic review filters combined with the child filter ranged from 68% to 96%; sensitivity of the child filter separately was 98%. The number of records retrieved with PubMed (limited to the years 1990-2006) by the systematic review filters combined with the child filter ranged from 7,861 to 618,053. Precision for the combined filters ranged from 2% to 52%. Due to poor reporting of specific systematic review criteria in both titles and abstracts, in 25% of the records screened we were 'unsure' if the article concerned a systematic review according to our definition.

Conclusions The high numbers of records yielded by sensitive search strategies and the low precision threaten the use of systematic reviews for clinical decision making and guideline development. Reporting of specific systematic review criteria in titles and abstracts is poor and reporting recommendations given by QUOROM should be used more strictly. To make identification using PubMed easier, there is an urgent need to set 'minimal' criteria that any review should fulfill in order for it to be indexed as a systematic review.

Introduction

Well conducted systematic reviews are seen as providing the best evidence to guide clinical practice, are cornerstones for the recommendations of evidence-based practice guidelines, and should be an integral part to the planning of future research.[1;2] In contrast to traditional narrative reviews, systematic reviews of the literature address a well-defined question, use an explicit search strategy to locate all relevant evidence, evaluate the retrieved studies using prospectively defined methodological criteria, and formally synthesize the results.[3;4]

Both clinicians and researchers should be able to reliably and quickly find systematic reviews. The Cochrane Library is the prime source of systematic reviews of the effects of healthcare interventions, but it does not contain all systematic reviews published in journals. Bibliographic databases such as MEDLINE do index Cochrane reviews, and can be used to identify other systematic reviews but only those which are indexed in the respective database. Yet, finding systematic reviews in MEDLINE poses two challenges. First, only a fraction of all citations in MEDLINE are for systematic reviews. Second, MEDLINE's indexing procedures do not include 'systematic review' as a 'Publication Type'. To limit the search results from a query in MEDLINE, it is therefore recommended that a methodological filter be used, consisting of text words and MeSH headings directed to general indicators of systematic reviews in the MEDLINE record.

Several search strategies for locating systematic reviews in MEDLINE have been developed and validated,[5-8] but their performance has never been evaluated in the universe of all articles included in MEDLINE. In addition, these filters have never been tested for their sensitivity and precision in finding child health systematic reviews.

The objective of the current study is to assess the usefulness of existing search filters in finding child health systematic reviews in MEDLINE using the PubMed interface by applying them to a reference standard of child health systematic reviews and by determining whether these filters focus the search strategy enough to be practical in the universe of MEDLINE. To this aim, the proportion of child health systematic reviews retrieved from our reference standard (sensitivity) and the number of 'true' child health systematic reviews from the searches' yield (precision) was calculated.

Methods

Inclusion and Exclusion criteria

In this article, we use the term systematic review to refer to any literature review, meta-analysis, or other article that explicitly indicates the use of a strategy for locating evidence by mentioning at least the databases that were searched, and reviewing the empirical evidence on children (<18 years of age). Excluded were systematic reviews assessing the effects of interventions in pregnant women on the fetus.

Existing search strategies for systematic reviews and child studies

To identify articles reporting on the development and validation of systematic review search filters in MEDLINE, we searched MEDLINE for the years 1995 to January 2006 with the following Medical Subject Heading (MeSH) terms: 'MEDLINE', 'Information Storage and Retrieval/methods' and 'Review, literature'. In addition, reference lists of relevant articles were reviewed and content experts were contacted to find additional studies. To improve precision, we combined the systematic review filters with a sensitive child filter, developed by the Cochrane Child Health Field (www.cochranechild health.org) to retrieve only the studies in children (Appendix 2).

Reference standard

To test the sensitivity of the previously mentioned search strategies, a reference standard set of systematic reviews was established by searching for child health systematic reviews in the Database of Abstracts of Reviews of Effects (DARE)[9] and by hand searching several pediatric journals for systematic reviews. All titles and abstracts in DARE (The Cochrane Library, Issue 2, 2004) were screened to find child health systematic reviews also indexed in MEDLINE. We hand searched seven pediatric journals with a range of Impact factors (Table 1). All issues of each journal were searched for five years: 1994, 1997, 2000, 2002, 2004. Additionally, we were interested if DARE contained all child health systematic reviews found by our hand search of pediatric journals. As DARE is only including systematic reviews on prevention, intervention or diagnosis, we selected systematic reviews covering these domains. Next, we searched DARE in July 2006 using The Cochrane Library interface to see how many child health systematic reviews on prevention, intervention or diagnosis found by our hand search were also included in DARE.

Sensitivity

As the child filter has not been validated yet, we calculated the sensitivity of the child filter separately. The sensitivity of a search strategy was defined as the proportion of child health systematic reviews retrieved from our reference standard set of child health systematic reviews.

$$\text{Sensitivity} = \frac{\text{Number of systematic reviews retrieved from reference standard}}{\text{Total number of systematic reviews in reference standard}} \times 100$$

Sensitivity was calculated for all systematic review filters separately and in combination with the child filter.

Precision

The precision of a combined search strategy was defined as the proportion of 'true' child health systematic reviews found in MEDLINE (limits 1990-2006) divided by the total number of records found.

$$\text{Precision} = \frac{\text{Number of 'true' systematic reviews}}{\text{All records in MEDLINE retrieved by the search}} \times 100$$

To calculate precision we selected a random sample of 100 records distributed evenly across each journal volume for the yield of each combined search strategy. Each sample was screened for 'true' systematic reviews. We regarded an article as a 'true' systematic review only if the title, abstract, author-supplied keywords or Publication Type terms explicitly identified the article as a systematic review or meta-analysis or if the article abstract indicated a strategy for locating the literature reviewed, and included children. We classified records as 'unsure' if: the title or abstract did not specifically state if children were included; it was stated that 'a literature search was performed', but without stating which databases were searched; or, if it was stated that the article was 'an evidence-based overview'. In all these cases full text was retrieved to make a final decision. Next, we applied the search strategies to seven *priority topics* in child health, developed by the Cochrane Child Health Field, that need systematic reviews. The number of records retrieved for the most sensitive and the most precise systematic review search strategies is reported and the precision for the most precise filter. The search strings for each clinical topic were created by using the Medical Subject Heading (MeSH) browser function.[10] We reported the time it takes a pediatrician (N.B.) and a clinical librarian (L.T.) to screen 100 MEDLINE records.

Results

Existing search strategies for identifying systematic reviews

We identified 4 studies reporting on the development and validation of several systematic review search strategies for MEDLINE.[5-8] We evaluated eight of the reported search strategies with varying sensitivity and precision and the PubMed Clinical Queries systematic review filter.[11] Appendix 1 shows the systematic review search strategies translated into PubMed format where necessary, and Appendix 2 the child search strategy we used.

Reference standard

Our reference standard contained 387 child health systematic reviews indexed in MEDLINE (Table 1). A total of 431 child health systematic reviews were identified in DARE (The Cochrane Library, Issue 2, 2004: n=4645), of which only 298 were indexed in MEDLINE. Hand searching MEDLINE-indexed pediatric journals identified 115 child health systematic reviews, of which 26 records were also included in DARE up till Issue 2, 2004. Of our hand searched 115 reviews, 73 were on prevention, intervention or diagnosis. By searching DARE in 2006, 38 of these 73 (52%) systematic reviews on prevention, intervention or diagnosis were also indexed in DARE.

Table 1 Reference standard: Number of systematic reviews found by hand searching seven pediatric journals with a range of impact factors, for the years 1994, 1997, 2000, 2002 and 2004 and number of child health systematic reviews included in DARE and indexed in MEDLINE.

Journal	Impact Factor	Systematic reviews/ meta-analysis (N)
Pediatrics	3.4	36
Journal of Pediatrics	3.2	6
Pediatric Infectious Diseases Journal	2.4	10
Archives of Diseases in Childhood	2.1	34
Archives of Pediatric and Adolescent Medicine	2.1	20
European Journal of Pediatrics	1.2	5
Journal of Perinatal Medicine	0.9	4
Total found by hand searching		115
DARE*		298
Overlap†		26
Total reference standard		387

* The Cochrane Library, Issue 2, 2004; reviews also indexed in MEDLINE

† Records found in both DARE up till Issue 2, 2004 and one of the journals hand searched. In 2006, 52% of the 'journal reviews' on prevention/ intervention or diagnosis were also indexed in DARE

Sensitivity

Sensitivity of the child filter when run separately in MEDLINE and tested against our reference standard was 98% (380 of 387; 95% CI: 96-99%). Sensitivity of the systematic review filters ranged from 68% to 96% (Table 2).

Precision

Not limited to a clinical topic

Figure 1 shows the number of records retrieved with the systematic review filters in PubMed (1990-2006) with and without the child filter added. The number of records retrieved with the systematic review filters ranged from 38,338 (Montori2) to 3,085,463 (Montori1) and for the combined filters from 7,861 (Montori2) to 618,053 (Montori1).

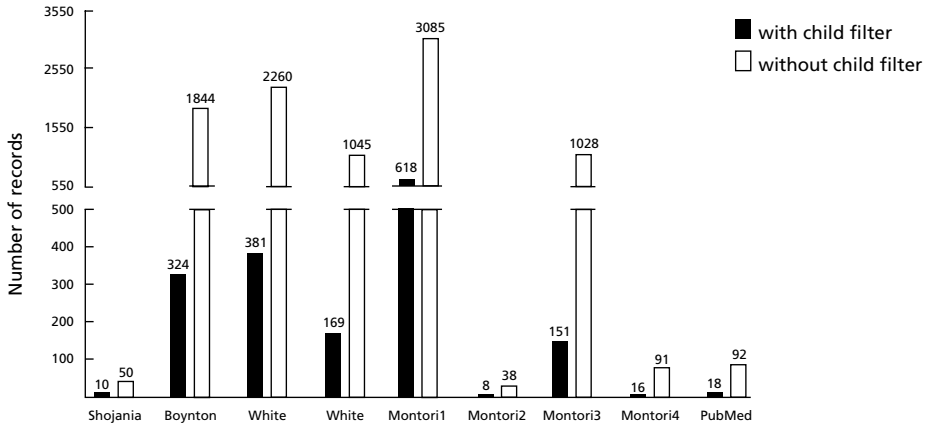
We screened 600 titles and abstracts for systematic reviews (a random sample of 100 records retrieved with 6 systematic review filters combined with the child filter, see Table 2) and were 'unsure' in 150 records whether the article was a 'true' systematic review. This generated 137 'true' systematic reviews and we were 'unsure' in 15 of those whether children were included in the review. Precision ranged from 2% to 52% in our study with the combined systematic review and child filters, compared to 3% to 69% reported in the original studies (Table 2). On average, it took the pediatrician 34 minutes to screen the titles and abstracts of 100 retrieved records and the clinical librarian 47 minutes.

Table 2 Reported performance and performance in our reference standard of combined systematic review filter with child filter

Search filters	Reported performance*		Performance in our reference standard	Performance In PubMed
	Sensitivity %	*Precision %*	*Sensitivity % (95% CI)*	*Precision % (95% CI)*‡
Shojania AND Child	93	50†	74 (69-78)	45 (36-55)
Boynton AND Child	98	20	95 (92-97)	3 (1-9)
White 1 AND Child	100	4	93 (91-96)	-
White 2 AND Child	94	10	94 (91-96)	2 (1-7)
Montori 1 AND Child	100	3	96 (93-97)	-
Montori 2 AND Child	71	57	68 (64-73)	52 (42-62)
Montori 3 AND Child	98	14	94 (91-96)	3 (1-9)
Montori 4 AND Child	74	69	72 (67-76)	-
PubMed AND Child	Unknown	Unknown	76 (72-80)	32 (24-42)

* Reported performance without child filter. Sensitivity child filter: 98% (95% CI: 96-99%). All numbers have been rounded.
† Applied to three clinical topics: screening for colorectal cancer, thrombolytic therapy for venous thromboembolism, and treatment of dementia
‡ Precision not calculated (-) for White1, Montori1 and Montori4 as other filters with the same sensitivity have lower number of hits.

Figure 1 Number of records retrieved through PubMed (1990-2006) for the systematic review filters with and without the child filter

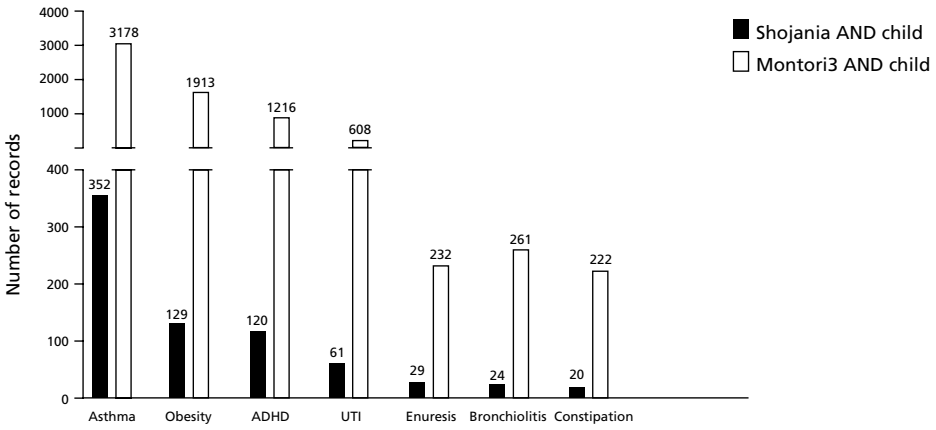


All numbers have been rounded

Limited to clinical topics

Figure 2 shows the number of records retrieved when applying a sensitive and a precise systematic review filter, in combination with the child filter, to 7 different clinical topics. Adding a MeSH term for asthma yielded the largest number of records and adding a MeSH term for constipation the lowest number (Shojania: 352 vs. 20; Montori3: 3178 vs. 222). Precision with the combined Shojania and child filter ranged from 49% for obesity and 93% for enuresis. On average, it took the pediatrician 18 minutes to screen the titles and abstracts of 100 topic-specific records for systematic reviews.

Figure 2 Number of records retrieved when applying a precise (Shojania) and a sensitive (Montori3) systematic review filter (combined with the child filter) in PubMed (1990-2006) to clinical topics



Abbreviations: ADHD = Attention Deficit Hyperactivity Disorder; UTI = Urinary Tract Infection

Discussion

There are a number of different types of systematic reviews to be identified in current bibliographic databases: these include Cochrane systematic reviews, which can be found in The Cochrane Library and systematic reviews published in journals ('journal reviews'). Journal reviews can also be found in The Cochrane Library's Database of Abstracts of Reviews of Effects (DARE). The DARE database provides quality assessments of systematic reviews published in journals identified by regular searches of important bibliographic databases such as MEDLINE, EMBASE, and CINAHL and by hand searching some of the major general journals, but not any specific pediatric journals. We have shown that DARE includes only 52% of our hand searched pediatric journal reviews. However, DARE uses more stringent inclusion criteria for systematic reviews than we did. The CDSR and DARE are already 'pre-filtered' in their focus on systematic reviews, so the only task is to find topic relevant systematic reviews in these databases. In general medical databases, like MEDLINE, systematic reviews should ideally be indexed using an appropriate and specific publication type term. Until there is an agreed definition of a systematic review, indexers will not be able to introduce or apply such a term. Pending this, we will need more precise search strategies to find systematic reviews in MEDLINE in order to reduce the number of irrelevant records.

The usefulness of a systematic review filter is expressed by sensitivity and precision. The trade-off between these two features will drive the choice of a filter. Researchers conducting a new systematic review or guideline developers would best be served by

the most sensitive search. This search will have the highest probability of retrieving all relevant reviews, but will have low precision, retrieving many irrelevant articles. Those with little time on their hands, for example clinicians looking for answers to patient care questions, will likely be best served by a more precise search strategy. The sensitivities of the systematic review filters as reported in the original studies and compared to our reference standard were quiet similar, except for the Shojania filter (93% vs. 74%, respectively). Shojania's reference standard mainly included 'high quality' systematic reviews, with probably higher standards of study description in the methods section of the abstract. A filter developed in such a sample will probably consist of another combination of text words and MeSH headings than would be optimal to retrieve systematic reviews in our reference standard.

As the sensitivity of the child filter alone was excellent (98%), we combined the systematic review filters with the child filter to improve precision. As before 1990 very few systematic reviews were published, we limited our search to 1990 and after. The number of records retrieved with the different systematic review filters was in the (hundreds of) thousands. Adding the child filter reduced the number of records by a factor of 5 to 7. Of the sensitive search filters, Montori3 has the best trade-off between sensitivity (94%), number of records retrieved (151,227) and precision (3%). Of the more precise filters, Shojania has the best trade-off between sensitivity (74%), number of records retrieved (10,188) and precision (45%). Although we added a child filter to improve precision and used a broad definition of a systematic review, precision was lower than reported in the original studies for which we have comparative precision data. The reason is probably that most existing search strategies were developed from a small, selective subset of systematic reviews published in 'high impact' journals. Precision may have been overestimated as it was not calculated in the universe of MEDLINE, but also from their subset of journals. These 'high impact' journals are more likely to publish systematic reviews. Shojania calculated precision in MEDLINE for three clinical topics (Table 2). As we have shown, precision varies considerably among different clinical topics and appeared to be higher than precision not limited to a clinical topic. Therefore, the results will not be generalizable.

It was often difficult to tell from the title and abstract if the article was a 'true' systematic review. Abstracts often lacked description of important criteria for systematic reviews. Instead, general terms were used like: 'a literature review was performed', 'we tried to collect all published literature', 'evidence-based', 'we performed a qualitative assessment of the literature', etc. Therefore, it is to be expected that the most sensitive systematic review filters also retrieve a lot of irrelevant records. In 1999 the Quality of Reporting Meta-analyses (QUOROM) statement was published to improve the quality of reporting meta-analyses or systematic reviews of clinical randomized controlled trials.[12] In order to be able to distinguish narrative reviews from systematic reviews, we advise that authors follow these recommendations when writing their abstract. In

addition, when children are included, this should be explicitly reported in the title or abstract. To make identification using PubMed easier, there is an urgent need to set 'minimal' criteria that any review should fulfill in order for it to be indexed as a systematic review. Ideally, a database of child health systematic reviews should be created.

Limitations

We used a broad definition of a systematic review and a number of these reviews will not fulfill some of the stringent criteria for systematic reviews used by others. Given the amount of time needed to perform a systematic review and the methodological skills required, our aim was not to exclude potentially relevant systematic reviews. The explicit mention of a strategy for locating evidence seems to be the most basic and least controversial.[7]

Our reference standard contains a subset of the MEDLINE database, as it is not possible to hand search all journals indexed in MEDLINE for child health systematic reviews, and could have over- or underestimated the sensitivity. We tried to compose a representative reference standard of systematic reviews by hand searching seven pediatric journals with a range of impact factors. This should help avoid selection bias in our search filter accuracy study because of higher standards of study description (e.g. by their use of structured abstracts with explicit methods section) in high-impact factor journals. In order to try to avoid possible biases occurring in any one year we sampled records from various different years within a ten years' range.

All systematic review filters, except Shojania's filter, were developed and tested using the OVID interface. Because PubMed is the only MEDLINE interface available free of charge worldwide and is widely used, we converted the OVID search filters into PubMed format.[5;6;8] Translating a search filter from OVID format to PubMed format is a factor that may influence the performance of the filter.

Recommendations for clinicians and researchers

We advise busy clinicians searching for systematic reviews on health care interventions to search The Cochrane Library (CDSR and DARE) for high quality systematic reviews. Several studies have shown that Cochrane reviews are more rigorous than journal reviews.[13;14] If no relevant systematic reviews are found in The Cochrane Library or the interest is other than prevention, treatment or diagnosis, other databases such as MEDLINE should be searched next. For a specific search Shojania's filter combined with the child filter can be used. For topics where not many reviews are available (e.g. constipation) it may be more efficient to search with the sensitive Montori3 filter combined with the child filter. Researchers and guideline developers interested in finding (close to) all systematic reviews can use Montori3 combined with the child filter. Readers can save the search filters they wish to use in PubMed and use them in combination with a clinical topic.

Acknowledgements

We are grateful to Marjan Loep †, Marcel van der Paardt and Arnold Leenders for searching DARE and hand searching the pediatric journals, Grace Liang for developing a program to enable random sample selection and Liza Bialy for screening records for 'true' systematic reviews.

Appendix 1

Systematic review search strategies

Shojania

((meta-analysis [pt] OR meta-analysis [tw] OR metanalysis [tw]) OR ((review [pt] OR guideline [pt] OR consensus [ti] OR guideline* [ti] OR literature [ti] OR overview [ti] OR review [ti]) AND ((Cochrane [tw] OR Medline [tw] OR CINAHL [tw] OR (National [tw] AND Library [tw])) OR (handsearch* [tw] OR search* [tw] OR searching [tw]) AND (hand [tw] OR manual [tw] OR electronic [tw] OR bibliographi* [tw] OR database* OR (Cochrane [tw] OR Medline [tw] OR CINAHL [tw] OR (National [tw] AND Library [tw]))))) OR ((synthesis [ti] OR overview [ti] OR review [ti] OR survey [ti]) AND (systematic [ti] OR critical [ti] OR methodologic [ti] OR quantitative [ti] OR qualitative [ti] OR literature [ti] OR evidence [ti] OR evidence-based [ti]))) BUTNOT (case* [ti] OR report [ti] OR editorial [pt] OR comment [pt] OR letter [pt])

Boynton (most sensitive strategy)

((Meta[tiab] NOT meta[ti]) OR (synthesis[tiab] NOT synthesis[ti]) OR (literature[tiab] NOT literature[ti]) OR (randomized[tw] NOT randomized[tiab]) OR published[tiab] OR Meta-Analysis[ptyp] OR extraction[tiab] OR (Trials[tw] NOT Trials[ti]) OR (controlled[tw] NOT controlled[tiab]) OR (search[tiab] NOT search[ti]) OR (medline[tiab] NOT medline[ti]) OR (selection[tiab] NOT selection[ti]) OR (sources[tiab] NOT sources[ti]) OR (review[tiab] NOT review[ti]) OR review[ptyp] OR articles[tiab] OR (reviewed[tiab] NOT reviewed[ti]) OR (english[tiab] NOT english[ti]) OR (language[tiab] NOT language[ti])) NOT (Letter[ptyp] OR comment[ptyp] OR editorial[ptyp])

White 1 (most sensitive strategy)

((Controlled[tiab] NOT controlled[ti]) OR (design[tiab] NOT design[ti]) OR (evidence [tiab] NOT evidence[ti]) OR (extraction[tiab] NOT extraction[ti]) OR "randomized controlled trials"[MeSH] OR Meta-Analysis[ptyp] OR Review[ptyp] OR (sources[tiab] NOT sources[ti]) OR (studies[tiab] NOT studies[ti])) NOT (Letter[ptyp] OR comment[ptyp] OR editorial[ptyp])

White 2 (most precise strategy)

((review[tiab] NOT review[ti]) OR Review[ptyp] OR meta-analysis[tiab] OR Meta-Analysis[ptyp]) NOT (Letter[ptyp] OR comment[ptyp] OR editorial[ptyp])

Montori 1 (most sensitive strategy)
search*[tiab] OR meta analysis[ptyp] OR meta-analysis[tiab] OR meta analysis[MeSH] OR review[ptyp] OR diagnosis[MeSH Subheading] OR associated[tiab]

Montori 2 (most precise strategy)
medline[tiab] OR (systematic[tiab] AND review[tiab]) OR meta-analysis[ptyp]

Montori 3 (minimising the difference between sensitivity and specificity)
Meta-analysis[ptyp] OR meta-analysis[tiab] OR meta-analysis[MeSH] OR review[ptyp] OR search*[tiab]

Montori 4 (combining most precise term with most sensitive terms)
Cochrane Database Syst Rev [ta] OR search[tiab] OR meta-analysis[ptyp] OR medline-[tiab] OR (systematic[tiab] AND review[tiab])

PubMed
(systematic review* [tiab] OR systematic literature review* OR meta-analysis [ptyp] OR meta-analysis [ti] OR metaanalysis [ti] OR meta-analyses [ti] OR evidence-based medicine OR (evidence-based AND (guideline [tiab] OR guidelines [tiab] OR recommendations)) OR (evidenced-based AND (guideline [tiab] OR guidelines [tiab] OR recommendation*)) OR consensus development conference [ptyp] OR health planning guidelines OR guideline[ptyp] OR cochrane database syst rev OR acp journal club OR health technol assess OR evid rep technol assess summ OR evid based dent OR evid based nurs OR evid based ment health OR clin evid) OR ((systematic [tiab] OR systematically OR critical [tiab] OR (study [tiab] AND selection [tiab]) OR (predetermined OR inclusion AND criteri*) OR exclusion criteri* OR "main outcome measures" OR "standard of care" OR "standards of care") AND (survey [tiab] OR surveys [tiab] OR overview* OR review [tiab] OR reviews [tiab] OR search* OR handsearch OR analysis [tiab] OR critique [tiab] OR appraisal OR (reduction AND risk AND (death OR recurrence))) AND (literature [tiab] OR articles [tiab] OR publications [tiab] OR publication [tiab] OR bibliography [tiab] OR bibliographies [tiab] OR published [tiab] OR unpublished OR citation OR citations OR database [tiab] OR internet [tiab] OR textbooks [tiab] OR references OR trials [tiab] OR meta-analysis [mh] OR (clinical [tiab] AND studies [tiab]) OR treatment outcome)) NOT (case report [ti] OR editorial [ti] OR editorial [ptyp] OR letter [ptyp] OR newspaper article [ptyp])

Appendix 2

Child search strategy

Infant[MeSH] OR Infant* OR infancy OR Newborn* OR Baby* OR Babies OR Neonat* OR Preterm* OR Prematur* OR Postmatur* OR Child[MeSH] OR Child* OR Schoolchild* OR School age* OR Preschool* OR Kid or kids OR Toddler* OR Adolescent[MeSH] OR Adoles* OR Teen* OR Boy* OR Girl* OR Minors[MeSH] OR Minors* OR Puberty[MeSH] OR Pubert* OR Pubescen* OR Prepubescen* OR Pediatrics[MeSH] OR Paediatric* OR Paediatric* OR Peadiatric* OR Schools[MeSH] OR Nursery school* OR Kindergar* OR Primary school* OR Secondary school* OR Elementary school* OR High school* OR Highschool*

Reference List

1 Haines A, Jones R. Implementing findings of research. BMJ 1994; 308(6942):1488-1492.

2 Scholten RJ, Kremer LC. [Systemic reviews as a basis for guidelines]. Ned Tijdschr Geneeskd 2004; 148(6):262-264.

3 Bero LA, Grilli R, Grimshaw JM, Harvey E, Oxman AD, Thomson MA. Closing the gap between research and practice: an overview of systematic reviews of interventions to promote the implementation of research findings. The Cochrane Effective Practice and Organization of Care Review Group. BMJ 1998; 317(7156):465-468.

4 Oxman AD, Cook DJ, Guyatt GH. Users' guides to the medical literature. VI. How to use an overview. Evidence-Based Medicine Working Group. JAMA 1994; 272(17):1367-1371.

5 Boynton J, Glanville J, McDaid D, Lefebvre C. Identifying systematic reviews in MEDLINE: developing an objective approach to search strategy design. Journal of Information Science 1998; 24:137-157.

6 Montori VM, Wilczynski NL, Morgan D, Haynes RB. Optimal search strategies for retrieving systematic reviews from Medline: analytical survey. BMJ 2005; 330(7482):68.

7 Shojania KG, Bero LA. Taking advantage of the explosion of systematic reviews: an efficient MEDLINE search strategy. Eff Clin Pract 2001; 4(4):157-162.

8 White V, Glanville J, Lefebvre C, Sheldon T. A statistical approach to designing search filters to find systematic reviews: objectivity enhances accuracy. Journal of Information Science 2001; 27:357-370.

9 Database of Abstracts of Reviews of Effects; Centre for Reviews and Disemination: available at: http://www.crd.york.ac.uk/crdweb/ [last accessed 17 December 2006].

10 Shaw RL, Booth A, Sutton AJ, Miller T, Smith JA, Young B et al. Finding qualitative research: an evaluation of search strategies. BMC Med Res Methodol 2004; 4:5.:5.

11 PubMed Clinical Queries - Systematic Reviews: available at: http://www.ncbi.nlm.nih.gov/entrez/query/static/clinical.shtml#reviews [last accessed 17 December 2006].

12 Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. Lancet 1999; 354(9193):1896-1900.

13 Jadad AR, Cook DJ, Jones A, Klassen TP, Tugwell P, Moher M et al. Methodology and reports of systematic reviews and meta-analyses: a comparison of Cochrane reviews with articles published in paper-based journals. JAMA 1998; 280(3):278-280.

14 Jadad AR, Moher M, Browman GP, Booker L, Sigouin C, Fuentes M et al. Systematic reviews and meta-analyses on treatment of asthma: critical evaluation. BMJ 2000; 320(7234):537-540.

CHAPTER 3

Neurodevelopment after neonatal hypoglycemia: a systematic review and design of an optimal future study

Nicole Boluyt, Anne A.M.W. van Kempen, Martin Offringa

Pediatrics 2006 Jun; 117(6): 2231-43

ABSTRACT

Objective Our goal was to assess the effect of episodes of neonatal hypoglycemia on subsequent neurodevelopment.

Methods We searched MEDLINE and EMBASE for cohort studies on subsequent neurodevelopment after episodes of hypoglycemia in the first week of life. Reference lists of available studies were reviewed and content experts were contacted for additional studies. Included studies were selected and appraised for methodological quality by two reviewers. Methodological quality was assessed according to well accepted criteria for prognostic studies. Eventually, all studies were given an overall quality score; poor, moderate, or high quality. Studies in the latter 2 categories were considered for quantitative data analysis.

Results Eighteen eligible studies were identified. The overall methodological quality of the included studies was considered poor in 16 studies and high in 2 studies. Pooling of results of the 2 high-quality studies was deemed inappropriate because of major clinical and methodological heterogeneity. None of the studies provided a valid estimate of the effect of neonatal hypoglycemia on neurodevelopment. Building on the strengths and weaknesses of existing studies we developed a proposal for an 'optimal' future study design.

Conclusions Recommendations for clinical practice cannot be based on valid scientific evidence in this field. To asses the effect of neonatal hypoglycemia on subsequent neurodevelopment, a well-designed prospective study should be undertaken. We submit a design for a study that may answer the still-open questions.

Introduction

Glucose is the essential substrate for brain function. Although important at all ages, it is particularly so in childhood since a normal supply is necessary to protect neural development.[1] Hypoglycemia is the most common metabolic problem in neonatal medicine.[2] Still, there is much controversy about the definition of a 'safe' blood glucose concentration (i.e. a value above which there is no risk of long-term neurodevelopmental impairment).

In 2000, a group of investigators in the field of hypoglycemia provided a consensus statement for the operational thresholds of blood glucose concentrations in the neonate.[3] They deliberately agreed on a high operational threshold (i.e. a 'safe' plasma glucose concentration [>2.5 mmol/L] that is applicable to any infant whether term or preterm). However, although several review articles on this topic have emerged,[1;2;4;5] no systematic review of the available studies on the prognosis after neonatal hypoglycemia exists. Also, existing reviews generally conclude that well-designed studies are scarce and that 'more research is needed' but fail to provide suggestions for an optimal design of these desired studies.

Our aim was to answer the following question: What is the neurodevelopmental outcome after neonatal hypoglycemia in the first week of life? To this end, we set out to find all relevant empirical studies on the prognosis after neonatal hypoglycemia in humans, appraise them for methodological quality, and summarize their results in a way that informs both clinical practice and subsequent research. The focus is on the available evidence with regard to the longer-term prognosis, that is, neurodevelopment, late neurophysiology or neuroradiology. Finally, building on the strengths and weaknesses of existing studies, we developed a proposal for an 'optimal' future study design.

Methods

Search Strategy

The objective of the literature search was to identify all cohort studies in neonates reporting on the long-term prognosis of hypoglycemia. Studies were identified by sensitive computerized searches of MEDLINE (1966 to January 2005) and EMBASE (1980 to January 2005). In MEDLINE the following Medical Subject Headings (MeSH) terms were used: blood glucose, glucose, hypoglycemia, limited to newborn (0-1 month) and human. In EMBASE, the MeSH terms glucose blood level and hypoglycemia were used, limited to infant (to 1 year) and human. In addition, reference lists of all available articles were reviewed to identify additional citations not found in the computerized search. Content experts in the field of neonatal glucose metabolism were contacted to find additional studies.

Inclusion Criteria

To be eligible for inclusion in this review, a study had to meet all of the following 3 criteria: (1) It had prospectively or retrospectively studied neonates <1 week after birth with hypoglycemia using any definition and any assay method. Retrospective studies were only eligible for inclusion when a prospective protocol was used to collect patient data. (2) Hypoglycemic neonates were compared with a control group or symptomatic versus asymptomatic hypoglycemia or the studies should provide information about other risk factors for adverse outcome. (3) At least 1 of the following outcome measures had been used: neurodevelopment, neurophysiology or neuroradiology at $\geq$1 year of age. Studies written in English, French, German, or Dutch were eligible for inclusion. Excluded were case-reports, studies with <3 patients, and 'abstracts only'.

Study Selection

Two investigators (N.B. and A.v.K.) independently selected studies as being potentially relevant on the basis of the titles and abstracts. Potentially relevant citations were retrieved in full text and checked for the presence of our eligibility criteria independently by 2 reviewers (N.B. and M.O.). Simple interobserver agreement for the latter step was calculated. Differences between the reviewers over which studies should be included were resolved by consensus.

Data Abstraction

Once a study met the inclusion criteria, 2 members of the research team (N.B. and M.O.) independently abstracted data by using structured data-abstraction forms. We captured information on the language of the report, study population, definition of hypoglycemia, glucose-assay method, treatment given for hypoglycemia, additional risk factors for adverse neurological outcome, neurodevelopmental outcome measurement, duration of follow-up, and main results. Disagreements were resolved by consensus.

Methodological Validity of Included Studies

To determine the methodological validity of the selected studies, 2 investigators assessed the design and execution of each study. Validity was assessed according to criteria for prognostic studies described by the Evidence-Based Medicine Working Group[6] with additional items described by Altman.[7] We specified each methodological item adjusted to the subject of our review. To maximize the validity of these questions we used a pilot sample of 5 articles in order to test the interrater agreement and came up with a final list of important items:

1 *Sample of patients*: A sample was considered to be 'well defined' if inclusion criteria, clinical characteristics, a definition of hypoglycemia, and when glucose was measured were sufficiently described. It was also assessed whether the study patients

were representative and whether they 'entered the cohort at a similar well-defined point in the course of the disease' (i.e. within 1 week after birth).

2 *Follow-up*: For neurodevelopment, neurophysiology and for neuroradiology a minimal follow-up period of 1 year was considered to be 'adequate'. Follow-up was considered to be 'complete' if the outcome was assessed in at least 80% of the study patients.

3 *Outcome*: It was assessed whether 'objective and unbiased outcome criteria' were applied. We considered outcome criteria to be 'objective' if there were predefined criteria for abnormalities (e.g. the Bayley Scales of Infant Development) and 'not objective' if there were no predefined criteria for abnormalities (e.g. neurological examination). Unbiased outcome assessment was defined as outcome measured by an observer who was masked for the status of all neonatal and possible other risk factors.

4 *Prognostic variables*: It was assessed whether other important factors associated with neurodevelopmental prognosis were assessed and whether there was adequate adjustment for prognostic factors in the data analyses. The most important prognostic factors for adverse neurodevelopment considered were (a) gestational age, (b) birth weight, (c) asphyxia, and (d) cerebral hemorrhage in preterm neonates <32 weeks' gestation. Adjustment was considered to be adequate in case of multivariate analyses of risk factors, exclusion of neonates with these prognostic risk factors for impaired neurodevelopment, or the use of controls matched for the 4 prognostic factors. Adjustment was considered inadequate when only descriptive data were given.

5 *Treatment subsequent to inclusion in cohort*: rated '+' if neonatal treatment for hypoglycemia was fully described and rated '-' if treatment was partially described or not described.

Studies were considered to be of (1) high methodological quality if the study group was well defined, follow-up was >80%, or if dropouts were described and shown to be comparable to the root population, objective outcome measures were used or, if not using objective outcome measures, outcome assessment should have been blinded, and if all important risk factors were assessed with adequate adjustment; of (2) moderate methodological quality if the study group was not well defined (-), follow-up was at least 50%, no objective outcome measures were used and no blinding of outcome assessment, and 3 of the 4 important risk factors were assessed with adequate adjustment; and of (3) poor methodological quality if <3 important risk factors were assessed

with descriptive data only or follow-up was < 50%. Only the studies judged to be of high or moderate methodological quality were considered for quantitative data analysis and are described in detail below.

A Priori Hypothesis Regarding Sources of Heterogeneity

Sources of heterogeneity in this systematic review were considered to be differences in study population, definition of hypoglycemia, types of outcomes assessed, follow-up time, and assessment and adjustment for other prognostic factors for adverse neurodevelopment.

Statistical Analysis

Pooling of data from individual studies to yield summary statistics of cumulative incidence of adverse outcome and relative risk measures for the various different prognostic factors was considered to be adequate in the absence of heterogeneity only.

Results

Study Selection

A total of 4508 references were screened in PubMed and 717 in EMBASE. We identified 45 potentially relevant articles.[8-52] Of the 45 potentially relevant articles on outcome, 17 were included[8;16-18;22;24;25;27;28;32;37;38;40;45;46;50;52] and 28 were excluded. Of the latter studies, 14 did not describe a control group and did not report other diseases or co-morbidity[10;13;14;19-21;23;29-31;34;35;47;49], 4 were retrospective without a prospective protocol,[11;33;41;42] 4 were cross-sectional studies,[15;36;44;51] 2 were case reports,[9;12] 1 was abstract-only,[39] 1 did not study neonates,[43] 1 outcome was brain section[26] and 1 study had a follow-up of only one week for neurodevelopment.[48] Checking reference lists did not yield additional studies. Contacting content experts in this field yielded 1 additional study that met our inclusion criteria, but it had not been published at the time of this review.[53] Thus, a total of 18 studies was included in our systematic review. The interobserver agreement was 87% for this selection. All differences could be resolved by consensus.

Description of the Selected Studies

Details on included studies are summarized in Table 1. Eighteen studies comprising 1583 infants were identified.[8;16-18;22;24;25;27;28;32;37;38;40;45;46;50;52;53] Three studies were clearly prospective studies by design,[28;32;46] 1 study was a combined prospective and retrospective design,[8] in 10 studies it was not clear whether the design was prospective or retrospective,[16;22;24;25;27;37;38;45;50;52] and 4 studies were definitely retrospective but used a 'prospective data collection protocol'.[17;18;40;53]

Table 1 Characteristics of included studies of neonates with hypoglycemia

Author (Year)	Patients		Plasma Glucose				Outcome		
	No.	*Clinical Characteristics*	*Definition of hypoglycemia, mmol/L*	*Assay Method*	*When Measured*	*Levels, mmol/L*	*Measurement Instruments*	*Other Risk Factors Assessed**	*Follow-up, %, Time*
TERM INFANTS, LGA									
Brand[53] (2004)	75	LGA term neonates (retrospective study with prospective protocol)	<2.2 1 h after birth and <2.5 h afterward	Glucose oxidase	1,3 and 5 h after birth	2.42 (0.6- 5.8) 1 h after birth; 2.57 (1.0- 4.0) 3 h after birth; 2.47 (0.8- 3.9) 5 h after birth	DDS, Snijders Oomen Non-verbal Intelligence test, and Child Behavior Check List at 4 y	**A:** na **B:** na **C:** as **D:** na **E:** as	64 (dropouts comparable to root population) at 4 y
TERM INFANTS, AGA AND SGA									
Gentz[24] (1969)	18	18 neonates with hypoglycemia (12 symptomatic, 6 asymptomatic), 35-40 wk gestation, 12 SGA (not clear whether retrospective or prospective)	<1.1	Glucose oxidase	< 72 h after birth	ns	Neurodevelopment (Bühler-Hetzer), neurological examination and EEG at 5-24 mo	**A:** as **B:** as **C:** as **D:** na **E:** as	50 at 1 y
Kinnala[32] (1999)	19	19 symptomatic hypoglycemic neonates, 36-42 wk, excluded if infections; 18 healthy, term controls (prospective study)	≤2.5	Enzymatic method using glucose-6-phosphate dehydrogenase	ns	Cases: 1.4 ± 0.7; controls 3.5 ± 0.7	MRI, ultrasound at 2 mo corrected gestational age; neurodevelopment at 1 y	**A:** as **B:** as **C:** ns **D:** na **E:** as	94 at 5-12 mo (mean: 11 mo)
PREMATURE INFANTS ONLY									
Lucas[40] (1988)	661	Premature, <1850 g, survived first 48 h (retrospective study with prospective protocol)	3 categories <0.6, <1.6 and <2.6	Glucose oxidase	Strip every 6 h, first 48-72 h; weekly plasma samples; neonates in intensive care: daily plasma samples	<2.6: 433 (66.5%) <1.6: 186 (28.1%) <0.6: 65 (9.8%)	Neurodevelopment: Bayley scales and neurological examination at 18 mo	**A:** na **B:** as **C:** as **D:** as **E:** ns	92% at 1.5 y

Author (Year)	Patients		Plasma Glucose				Outcome		
	No.	*Clinical Characteristics*	*Definition of hypoglycemia, mmol/L*	*Assay Method*	*When Measured*	*Levels, mmol/L*	*Measurement Instruments*	*Other Risk Factors Assessed**	*Follow-up, %, Time*
PREMATURE AND TERM INFANTS									
Creery[16] (1966)	22	22 symptomatic hypoglycemic neonates, premature (n=2) and full term (n=20) (not clear whether retrospective or prospective)	≤1.1	Haslewood and Strookman	Infants showing abnormal behavior	ns	Mental retardation at 5 mo to 4 y	**A:** as **B:** as **C:** ns **D:** na **E:** as	88 at 1 y
Haworth[27] (1967)	46	23 symptomatic and 23 asymptomatic hypoglycemic neonates (not clear retrospective or prospective)	≤1.1	Glucose oxidase or true sugar method of Nelson and Somogyi	When symptoms appear	ns	Neurodevelopment at 0.5-4 y	**A:** as **B:** as **C:** ns **D:** ns **E:** ns	70 at 0.5-4 y
Griffiths[25] (1971)	41	41 hypoglycemic neonates, 41 controls, matched for birth weight, Apgar, gestational age and social class (not clear retrospective or prospective)	<1.1	Modification of the method of Watson	Majority on first day of life	ns	Neurodevelopment: IQ and locomotor quotient at 4 y	Controls matched for **A:** as **B:** as **C:** as **D:** na **E:** as	36 at 4 y
Kumari[38] (1971)	19	19 premature or term neonates with idiopathic transient symptomatic hypoglycemia (not clear retrospective or prospective)	<1.1 for birth weight ≤ 2.0 kg and <1.7 for term infants	Method of Astoor and King	At first suspicion of hypoglycemia	ns	Neurologic status at 3-18 mo	**A:** as **B:** as **C:** as **D:** na **E:** as	21 at 1 y
Koivisto[37] (1972)	151	151 neonates with exclusively hypoglycemia and 56 controls; 4 groups: symptomatic convulsion, symptomatic non-convulsion, asymptomatic, and controls (not clear retrospective or prospective)	<1.7	Glucose oxidase or Hultman method	Screening of ''high risk" group and symptomatic; 3-times-daily glucose measurements for 2-4 d	ns	Neurologic evaluation and developmental evaluation at 1-4 y	**A:** ns **B:** ns **C:** as **D:** as **E:** as	100 at 1-4 y

Author (Year)	Patients		Plasma Glucose				Outcome		
	No.	*Clinical Characteristics*	*Definition of hypoglycemia, mmol/L*	*Assay Method*	*When Measured*	*Levels, mmol/L*	*Measurement Instruments*	*Other Risk Factors Assessed**	*Follow-up, %, Time*
PREMATURE AND TERM INFANTS									
Pildes[46] (1974)	39	39 hypoglycemic term and preterm neonates; 41 controls matched by weight, gender, gestation, race, mode of delivery, and condition of mother and infant (prospective study)	≤1.7	Glucose oxidase	After 3-4 h fasting or at time of symptoms	ns	Neurologic examination, EEG, psychological testing at 5-7 y	Controls matched for **A:** as **B:** as **C:** as **D:** as **E:** as	25 at 5-7 y
Fluge[22] (1975)	67	50 LBW infants and 17 term infants with hypoglycemia; 3 groups: asymptomatic, symptomatic transient/ no other neonatal complications (asphyxia, brain injury, RDS), and secondary (unsatisfactory response to glucose infusion) (not clear whether retrospective or prospective)	ns	ns	Routine blood glucose determinations among 323 LBW infants; term: ns	ns	Development: questionnaire and neurologic status; EEG; age: 2.5–4.75 y	**A:** ns **B:** ns **C:** as **D:** as **E:** as	55 at 2.5-4.75 y
Abel[8] (1987)	58	58 hypoglycemic neonates, 1 of a twin (combined retrospective and prospective study)	ns	ns	ns	ns	Development according to Sälzler at 3 y	**A:** ns **B:** as **C:** ns **D:** ns **E:** ns	72 at 3 y
Singh[50] (1991)	72	72 hypoglycemic preterm and term neonates of 2248 selectively screened for hypoglycemia; excluded were neonates with associated risk factors for adverse neurodevelopment (not clear whether retrospective or prospective)	<1.7	Dextrostix, verified by glucose oxidase method	Dextrostix 2,4,8, 12,24 h after birth or symptoms; if hypoglycemic: hourly blood sugar estimations	Symptomatic: 0.7 (± 0.3); asymptomatic 1.3 (± 0.2)	Neurodevelopment: Bayley scales and neurologic examination: at 1 y	**A:** ns **B:** ns **C:** as **D:** as **E:** as	81 at 1 y
Yamaguchi[52] (1997)	135	1561 high-risk neonates, of which 135 were hypoglycemic (not clear whether retrospective or prospective)	Term: <1.9; preterm: <1.4	ns	ns	ns	Neurodevelopment at 2, 4 and 6 y: developmental quotient and IQ for VLBW infants only	**A:** ns **B:** as **C:** as **D:** ns **E:** ns	ns

Author (Year)	Patients		Plasma Glucose				Outcome		
	No.	*Clinical Characteristics*	*Definition of hypoglycemia, mmol/L*	*Assay Method*	*When Measured*	*Levels, mmol/L*	*Measurement Instruments*	*Other Risk Factors Assessed**	*Follow-up, %, Time*
INFANTS OF MOTHERS WITH DIABETES									
Haworth[28] (1976)	25	Infants of diabetic mothers, 35-38 wk gestation, 25 hypoglycemic and 12 nonhypoglycemic (prospective study)	≤1.1 for birth weight <2.5 kg; ≤1.7 for birth weight ≥2.5 kg	Cord blood: according to the method of Huggett and Nixon; capillary blood: ns	Cord blood and 1,2,3,6,12,24,48, 72 h after birth	Hypoglycemic: lowest glucose 0.6, nonhypoglycemic 2.0; No. of episodes and duration of hypoglycemia also mentioned	Neurologic examination and developmental evaluation (Yale Developmental Schedules) at 3 y	**A:** as **B:** as **C:** as **D:** na **E:** as	75 at 3 y
Persson[45] (1984)	94	64 neonates of mothers with insulin-dependent diabetes mellitus and 21 neonates of mothers with gestational diabetes; gestation ≥28 weeks; only 18 had glucose <1.7 mmol/l (not clear whether retrospective or prospective, pobably retrospective)	<1.7	ns	2, 4,12, 14,48 h after birth	ns	Development: interviews with mothers and neurologic examination; IQ (Terman-Merril method) at 5 y	**A:** as **B:** as **C:** as **D:** ns **E:** as	83 at 5 y
PERSISTEND HYPERINSULINEMIC HYPOGLYCEMIA OF INFANCY									
Cresto[17] (1998)	26	Persistent hyperinsulinemic hypoglycemia of Infancy (PHHI) treated with diazoxide or diazoxide followed by surgery (retrospective study with prospective protocol)	ns	Glucose oxidase	ns	Insulin-glucose ratio	EEG, intellectual performance, school performance, age not mentioned	**A:** ns **B:** ns **C:** ns **D:** ns **E:** ns	ns
Dacou-Voutetakis[18] (1998)	15	15 infants with persistent hyperinsulinemic hypoglycemia of Infancy (retrospective study with prospective protocol)	ns	ns	ns	ns	Neurologic status (school progress and social behavior) at 2–20 y	**A:** ns **B:** ns **C:** ns **D:** ns **E:** ns	87 at 2-20 y

Abbreviations: na indicates not applicable (e.g. gestation in only term neonates, birth weight in only LGA neonates, or cerebral hemorrhage in term neonates); as, assessed; ns, not stated; LGA, large for gestational age; SGA, small for gestational age; LBW, low birth weight; EEG, electroencephalogram; VLBW, very low birth weight

* Risk factors: A, gestation; B, birth weight; C, asphyxia; D, cerebral hemorrhage; E, other (e.g. sepsis, use of antibiotics, maternal history, etc.)

Three studies included term infants only; in 1 study, large-for-gestational-age (LGA) term infants (total n=75) were included,[53] and 2 studies included small-for-gestational-age (SGA) and appropriate-for-gestational-age (AGA) term infants (total n=37).[24;32] One study included premature infants only (total n=661),[40] and 10 studies included both premature and term infants (total n=650).[8;16;22;25;27;37;38;46;50;52] Two studies included infants of mothers with diabetes (total n=119),[28;45] and 2 studies included neonates with persistent hyperinsulinemic hypoglycemia (total n=41).[17;18] Four prevalence and follow-up studies described all at-risk infants who were included.[28;40;45;53] Fourteen studies described only hypoglycemic infants,[8;16-18;22;24;25;27;32;37;38;46;50;52] of which 5 studies reported on symptomatic and asymptomatic hypoglycemic infants.[22;24;27;37;46]

Fourteen studies provided a definition of hypoglycemia: 2 studies with term infants[32;53] (<2.5 or $\leq$2.5 mmol/L), 7 studies with both premature and term infants[28;37;38;45;46;50;52] (<1.7 or $\leq$1.7-1.9 mmol/L), 6 studies with both premature and term infants[16;24;25;27;28;38] (<1.1 mmol/L), and 1 study using 3 categories[40] (<0.3, <1.6, and <2.6 mmol/L). Two studies used different definitions for different birth weights.

The assay method used to measure blood glucose was reported in 13 studies,[16;17;24;25;27;28;32;37;38;40;46;50;53] 8 of which used the glucose-oxidase method;[17;24;27;37;40;46;50;53] the other 5 used various other methods.[16;25;28;32;38] Thirteen studies reported what time the blood glucose was measured: routinely at preset times in 6 studies,[28;40;45;46;50;53] routinely but not at preset times in 3 studies,[22;24;37] and in case of signs and symptoms of hypoglycemia in 6 studies.[16;27;37;38;46;50] Two studies measured glucose both routinely and in case of symptoms.[46;50] Mean blood glucose levels were reported in 5 studies[28;32;40;50;53] and varied between <0.6 and 2.6 mmol/L.

Follow-up outcome variables that were used were (1) neurodevelopment, using various different scales, questionnaires, and neurological examination (all 18 studies[8;16;18;22;24;25;27;28;32;37;38;40;45;46;50;52;53]), (2) neuroradiology (1 study[32]), and (3) neurophysiology, in all cases using the electroencephalogram (4 studies[17;22;24;46]). Mean follow-up time ranged from 2 months to 20 years across studies. Completeness of follow-up ranged from 21%-100%. Seven studies assessed all 4 defined adverse-outcome risk factors.[24;25;28;38;40;46;53] Three studies assessed 3 risk factors,[16;32;45] 5 studies assessed 2 risk factors,[22;27;37;50;52] 1 study assessed 1 risk factor,[8] and 2 studies assessed none of the defined risk factors.[17;18]

Validity

Data on the validity of the studies are shown in Table 2. The study population was 'well defined', representative, and included in the study at a similar point in the disease in 7 of 18 studies.[25;28;37;40;45;46;53] Follow-up was complete in 8 of 18 studies.[16-18;32;37;40;45;50] The outcome variable for at least 1 outcome variable was 'objective' and blinded in 6 of 18 studies.[25;28;40;46;50;53] Four studies adequately adjusted for our defined other risk factors for impaired neurodevelopment.[25;40;46;53] Treatment of cases of neonatal hypo-

Table 2 Methodologic quality of included studies of neonates with fypoglycemia

Author (Year)	Study group[1]	Follow-up[2]		Outcome[3]	Other Risk Factors[4]	Treatment Described	Overall Quality
		Measurement Instruments	*Complete*				
TERM INFANTS, LGA							
Brand[53] (2004)	+	DDS	-	+	+	+	High
		SON-IQ test	-	+			
		Child Behavior Checklist	-	+			
TERM INFANTS, AGA and SGA							
Gentz[24] (1969)	-	Neurologic examination	-	-	-	+	Low
		Neurodevelopment	-	-			
		EEG	-	-			
Kinnala[32] (1999)	-	Neurodevelopment:	+	-	-	+	Low
		MRI, ultrasound	-	-			
PREMATURE INFANTS ONLY							
Lucas[40] (1988)	+	Bayley scales	+	+	+	+	High
		Neurologic examination	+	-			
PREMATURE AND TERM INFANTS							
Creery[16] (1966)	-	Mental retardation	+	-	-	+	Low
Haworth[27] (1967)	-	Neurodevelopment	-	-	-	+	Low
Griffiths[25] (1971)	+	Neurodevelopment: IQ and locomotor quotient	-	+	+	-	Low
Kumari[38] (1971)	-	Neurologic status	-	-	-	+	Low
Koivisto[37] (1972)	+	Neurologic evaluation	+	-	-	+	Low
		Developmental evaluation	+	-			

Author (Year)	Study group[1]	Follow-up[2]		Outcome[3]	Other Risk Factors[4]	Treatment Described	Overall Quality
		Measurement Instruments	*Complete*				
Pildes[46] (1974)	+	Neurologic examination	-	-	+	+	Low
		EEG	-	-			
		Psychological testing	-	+			
Fluge[22] (1975)	-	Development	-	-	-	-	Low
		EEG	-	-			
Abel[8](1987)	-	Development	-	-	-	-	Low
Singh[50] (1991)		Bayley scales	+	+	-	-	Low
		Neurologic examination	+	+			
Yamaguchi[52] (1997)	-	Developmental quotient	ns	-	-	-	Low
		IQ scores	ns	-			
INFANTS OF MOTHERS WITH DIABETES							
Haworth[28] (1976)	+	Neurologic examination	-	-	-	-	Low
		Developmental evaluation	-	+			
Persson[45] (1984)	+	Development	+	-	-	-	Low
		IQ (Terman-Merril method)	+	-			
PERSISTENT HYPERINSULINEMIC HYPOGLYCEMIA OF INFANCY							
Cresto[17] (1998)	-	EEG	+	-	-	-	Low
		Intellectual performance	+	-			
		School performance	+	-			
Dacou-Voutetakis[18] (1998)	-	Neurologic status (school progress and social behavior)	+	-	-	-	Low

See 'Methodologic Validity of Included Studies' (under 'Methods') for details. ns indicates not stated; SON-IQ test, Snijders Oomen nonverbal intelligence test; EEG, electroencephalogram; +, criterion met; -, criterion not met.

1 Well defined, representative and at similar point in disease.

2 Complete.

3 Objective and blind.

4 All other important risk factors associated with adverse neurodevelopment were assessed with adequate adjustment.

glycemia was fully described in 9 studies.[16;24;27;32;37;38;40;46;53] Only 2 studies were considered to be of high methodological quality,[40;53] none of moderate quality, and 16 (89%) of poor quality.

Effect of hypoglycemia on long-term prognosis

The main results of all included studies are listed in Table 3. Some studies found no differences between neonates with and without episodes of hypoglycemia on neurodevelopment, whereas others showed serious brain damage after episodes of neonatal hypoglycemia. Pooling of results was considered to be inadequate because of both major clinical heterogeneity between studies and methodological limitations in the majority of studies.

Table 3 Results of included studies of neonates with hypoglycemia

Author (Year)	Main results
TERM INFANTS, LGA	
Brand[53] (2004)	Hypoglycemia in 60 (80%) of neonates; no significant differences between neonates with and without hypoglycemia in DDS, SON-IQ and CBCL total scores; 1 IQ subscale (reasoning) scores statistically significant lower in hypoglycemic neonates: 107.6 (15.0) and 116.9 (16.1); no significant correlation between lowest plasma glucose level and SON-IQ
TERM INFANTS, AGA and SGA	
Gentz[24] (1969)	2 (17%) of 12 symptomatic neonates had serious brain damage (not defined), 1 (8%) of 12 had a borderline value for neurodevelopment and slight muscular hyper tonicity, 9 (75%) of 12 normal; 6 (100%) of 6 asymptomatic hypoglycemic neonates had normal results on all tests
Kinnala[32] (1999)	MRI, ultrasound: RR 3.7 (90% CI: 1.11 to 12.28) for any abnormality in the brain; good tendency for recovery (5 of 7); 17 (94%) of 18 cases normal neurodevelopment
PREMATURE INFANTS ONLY	
Lucas[40] (1988)	Maximum slope regression coefficient and significance for motor and mental development when a cutoff of 2.5 mmol/L was used; RR 3.5 (95% CI: 1.3 to 9.4) for cerebral palsy or developmental delay in infants with hypoglycemia on ≥5 d compared to those with no hypoglycemia
PREMATURE AND TERM INFANTS	
Creery[16] (1966)	5 (23%) of 22 died; 7 (32%) of 22 suffered brain damage (5 primary cerebral birth injury, 2 infantile spasms), 10 (45%) of 22 were normal
Haworth[27] (1967)	17 symptomatic: - 12 idiopathic (rapid relieve of symptoms after glucose administration): 3 (25%) of 12 with mental retardation - 2 secondary to cerebral disease: 1 (50%) of 2 was retarded - 3 insulin-dependent diabetes mellitus: 2 (67%) of 3 were retarded 15 asymptomatic: - 2 (13%) of 15 were possibly retarded
Griffiths[25] (1971)	Cerebral damage: 6 (14.6%) of 41 cases, 5 (12.2%) of 41 controls; IQ: 99.8±10.2 (cases), 100.4±11.9 (controls); locomotor quotient: 102.6±8.1 (cases), 102.3±9.3 (controls)

Kumari[38] (1971)	4 (21%) of 19 died, 1 (5%) of 19 had neurologic abnormalities, 11 (55%) of 19 were normal, 3 (16%) of 19 were lost to follow-up
Koivisto[37] (1972)	Symptomatic convulsion: 4 (50%) of 8 damaged (2 infantile spasms, 1 severe motor retardation, 1 developed idiopathic hypoglycemia later in life). Symptomatic nonconvulsion: 9 (12%) of 77 damaged Asymptomatic 4 (6%) of 66 pathologic Controls: 3 (5%) of 56 pathologic
Pildes[46] (1974)	After 2 y, 22 of 27 hypoglycemic neonates and 15 of 29 controls had abnormalities on neurologic examination; no significant differences in EEG abnormalities; at 1-4 y, no significant differences in mean IQ; at 5-7 y: 13 of 26 hypoglycemic infants and 6 of 27 controls had an IQ <86
Fluge[22] (1975)	Asymptomatic: 1 (14%) of 7 had MBD Symptomatic transient: 3 (33%) of 9 had MBD or severe retardation Secondary: 4 (19%) of 21 had MBD or developmental delay
Abel[8] (1987)	45% of hypoglycemic neonates has a developmental delay
Singh[50] (1991)	Symptomatic: MDI: 76.6±10.3, PDI : 74.5±13.1 Asymptomatic: MDI 93.2±5.7, PDI 94.0±5.5; controls: MDI 92.2±6.2, PDI 91.1±6.1 Duration of hypoglycemia was directly related to the neurodevelopmental outcome
Yamaguchi[52] (1997)	15 (11%) of 135 cases had major neurologic handicaps; DQ at 2,5 y 103.8±24.2 for cases vs 116.8±20.7 for controls; IQ at 6 y: 97.9±22.2 for cases vs 106.6±20.7 for controls; not significant
INFANTS OF MOTHERS WITH DIABETES	
Haworth[28] (1976)	Hypoglycemic: 7 (28%) of 25 were neurologically abnormal, DQ 94±2; nonhypoglycemic: 4 (33%) of 12 were neurologically abnormal, DQ 98±3
Persson[45] (1984)	18 neonates with glucose <1.7 mmol/L: all normal IQ; no relation between plasma glucose determined at 2-4 h after birth and IQ at follow-up
PERSISTENT HYPERINSULINEMIC HYPOGLYCEMIA OF INFANCY	
Cresto[17] (1998)	10 of 26 had neurologic abnormalities: 4 had brain damage, 4 had seizures, and 2 had slight motor disability; 4 had an IQ <60
Dacou-Voutetakis[18] (1998)	13 (100%) of 13 showed normal development

Abbreviations: DDS, Denver Developmental Scale; SON-IQ, Snijders Oomen Nonverbal Intelligence Test; CBCL, Child Behavior Checklist; RR, Relative Risk; DQ, Developmental quotient; MBD, minimal brain dysfunction; MDI, Mental Developmental Index; PDI, Psychomotor Developmental Index.

Description of the two high-quality studies

The study by Brand et al[53] evaluated the effects of transient hypoglycemia on the first postnatal day in 75 healthy term LGA infants on their neurodevelopmental outcome at 4 years of age. There were no significant differences between children with normoglycemia and hypoglycemia in Denver Developmental Scale (DDS) scores (total score, % definitively normal): 93% in 60 hypoglycemic children and 100% in 15 normoglycemic children (95% confidence interval [CI] of the difference: -14 to 16%) and Child Behavior Checklist scores (51.3 vs 52.9; 95% CI of the difference: -5.0 to 8.2%). Although total IQ did not differ significantly between hypoglycemic and normoglycemic children (108.9 vs 114.3; 95% CI of the difference: -2.8 to 13.5%), 1 IQ subscale (rea-

soning) did (107.6 vs 116.8, 95% CI of the difference: 1.3 to 17.2%). Because 19 items were tested, it is to be expected that 1 item shows a significant difference by chance. In addition, no relation was found between the lowest glucose level and the Snijders-Oomen Non-verbal Intelligence Test scores in individual subjects.

The study by Lucas et al[40] reported the incidence of hypoglycemia (i.e., plasma glucose concentration <2.6 mmol/L) in 661 premature infants and related the occurrence and persistence of hypoglycemia to the neurodevelopmental outcome at 18 months of age. In the 31 infants with hypoglycemia recorded on ≥5 separate days, the authors found reduced mental score (-14 points; 95% CI: -22 to -6) and motor score (-13 points; 95% CI: -20 to -5) on the Bayley developmental scale at 18 months of corrected age, compared with 177 nonhypoglycemic infants, even after adjustment for a wide range of factors known to influence development. Also, the incidence of cerebral palsy or developmental delay (i.e., mental or motor score of ≤70) was increased by a factor 3.5 (95% CI: 1.3 to 9.4). These data suggest that moderate hypoglycemia, if prolonged, is associated with an increased risk of impaired neurodevelopment.

Discussion

Hypoglycemia is the most common metabolic problem in neonates. Still, the level or duration of hypoglycemia that is harmful to an infant's developing brain is not known.

Our findings are the result of a systematic search for all relevant studies on the neurodevelopmental outcome after neonatal hypoglycemia and a critical appraisal of the methodological quality of these studies. The results of the individual studies are quite conflicting: some studies found no differences between neonates with and without episodes of hypoglycemia on neurodevelopment, whereas others show serious brain damage after episodes of neonatal hypoglycemia.

Several sources for this variability were identified. First, there is wide agreement that a reliable prognostic study requires a well-defined cohort of patients at the same stage of their disease course. In our review the majority of included studies failed to define the study population and the timing and number of hypoglycemic episodes adequately.

Second, a well known problem in retrospective studies is the fact that they largely depend on the memory of patients or their parents or on the accuracy and completeness of data in medical charts. In our review only 3 studies were clearly prospective in design and execution; in 55% of studies it was not clear whether it was a prospective or retrospective design.

Third, the accuracy of the measurement of blood glucose is another important source of variability. In our review, 72% of the studies reported which assay method was used, but only 44% of these studies used the reliable glucose-oxidase method.

Next, assessment of outcomes should be blinded for prognostic information. This is especially important for outcomes requiring a great deal of judgment, as is the case for some neurodevelopmental tests and neurologic examination. In our review, only 38% of the studies used blinded outcome assessment.

Then, follow-up should be long enough to detect the outcomes of interest. Several studies have shown that cognitive delay, particularly in the mildest forms, cannot be predicted accurately from tests in early infancy.[54] Ideally, in this field of research, an intelligence test should be performed at, for example, 7 years of age. In our review, only 11% of the studies had a follow-up of 5 to 7 years of age, and potential underestimation of the risk of impairment of cognitive performance may have occurred.

Next, to get a valid picture of the relative impact of the primary prognostic variable (i.e. a low neonatal blood sugar level), it is important to account for the presence of other prognostic variables that may confound the observed relationship with impaired neurodevelopment. Adjustment by multiple-regression analysis or by stratification according to the presence or absence of other prognostic factors for adverse neurodevelopment should be performed. Because the choice of prognostic factors in any analysis is arbitrary, we chose the 4 prognostic factors that should be addressed anyway in this field. We are aware that there are other prognostic factors that affect neurodevelopment, such as 'chronic lung disease' and 'time spent on the ventilator'. Still, only 22% of the studies assessed the defined 'minimum' of other prognostic factors for adverse neurodevelopment and adjusted the results accordingly. In this line of thought, the acute treatment of neonatal hypoglycemia is another important issue. If the infants' treatment varies with the presence of other prognostic variables, then a study cannot provide an unbiased and meaningful assessment of the prognostic significance of neonatal hypoglycemia episodes per se. Neonatal hypoglycemia treatment was fully described in only 50% of the included studies. Ideally, prognostic variables should be evaluated in a cohort of patients treated according to a protocol, such as in a randomized trial.

As in other varieties of retrospective empirical research, reviewing the medical literature is prone to several types of bias. The one most known is publication bias, a form of selection bias in systematic reviews in which 'positive' results have a better chance of being published and included than 'negative' results. This holds especially for intervention research, where a study showing beneficial treatment effect is more likely to be published than a study showing no treatment effect or an adverse effect. In our sample of prognostic studies, we found heterogeneous results, suggesting that publication bias is probably not a great problem. Second, it is possible that we missed some studies. However, a comprehensive search of the published literature was conducted and experts in this field were asked for additional studies. We therefore believe that no important studies were missed. Another potential limitation of our review is the fact that we only included studies written in English, French, German, or Dutch.

Several studies have shown that language restrictions do not chance the results of systematic reviews.[55;56] Another type of bias is information bias; the available information in the articles may be incomplete or may be generated by using methods that are of substandard quality. In this case, reviewers need to be cautious with both the validity of the information they include in their review and the completeness of the material. In our review, data were always abstracted independently by 2 investigators using structured data-abstraction forms. In addition, the methodological quality was also assessed by 2 investigators. We used a widely quoted and internationally used checklist for the methodological quality of prognostic studies described by the Evidence-Based Medicine Working Group, with some important additional items.[6;7] One difficulty with any quality-assessment tool is that answering the question for each item often requires judgment. In order to overcome this problem, we prespecified each general methodological question from the list for our review and used a pilot sample to test the interrater agreement. Using these context-specific questions it appeared that little discussion about differences in the answers to the methodological questions was needed. As a result, there was 100% agreement about the 'overall methodological quality' of the individual studies.

We conclude that the methodological quality of the majority of the included studies is poor. Even the two 'high-quality' studies have limitations. The study by Brand et al is a retrospective study, underpowered to detect differences in IQ smaller than 15 points. Also, the motivations for starting or withholding glucose treatment are unavailable. A minor point is the fact that the DDS that was used is less comprehensive and sensitive than the Bayley Scales of Infant Development. The study by Lucas et al was not a prospective study to investigate the long-term effects of neonatal hypoglycemia. Rather, retrospective data were obtained for a group of preterm infants with birth weights <1850 grams who had been enrolled in a multicenter feeding study. The investigators adjusted the results for cerebral hemorrhage, but only 2 of 5 participating centers conducted routine ultrasound investigations. At the time of the study in 1988, it was probably not possible to reliably detect small cerebral hemorrhages on ultrasound, which may, on the other hand, not be major confounders of the observed relationship between neonatal hypoglycemia and neurodevelopment. Furthermore, in an observational study design it is hard to prove that the association between modest hypoglycemia and poor neurodevelopment is causal or just reflects failure to adjust for all potential (known and unknown) confounders. The clinical importance of a reduction in mental and motor development scores of 14 and 13, respectively, at 18 months may be questioned. Although never formerly published in a paper, in his response to a 'letter to the editor', Lucas produced long-term follow-up results of this study: at 7.5 to 8 years, evidence of persisting associations between neonatal hypoglycemia and lower test scores were found in 2 of the 4 outcomes, arithmetic and motor tests, with ~ 0.5 SD reduction in scores (adjusted for respiratory support, birth weight, and gesta-

tion) after the neonatal concentration of blood glucose was <2.6 mmol/L for >3 days ($p<0.005$).[57]

We considered statistical pooling of the results of the included studies to get an estimate of the overall effect of episodes of neonatal hypoglycemia on subsequent neurodevelopment. Heterogeneity is a major threat to the validity of such meta-analyses and can be ascribed to differences in study methods, study populations, interventions, outcomes, or chance.[58] In our study we found, as stated earlier, major clinical heterogeneity in patient characteristics, definitions of hypoglycemia, assay methods, treatment protocols, length of follow-up, assessed outcomes, and methodological quality (Table 1). Therefore, we decided that the statistical pooling of results was inappropriate.

None of the studies that we reviewed could validly quantify the effect of episodes of neonatal hypoglycemia on subsequent neurodevelopment. In the last 15 years, several authors have urgently requested methodologically sound studies.[3;59-62] To our knowledge, however, no one has come up with a design for such a future study that overcomes the wide range of problems that we encountered, and as of yet, these studies have not been performed, probably because of the perceived complex nature of the study of hypoglycemia in the neonate. Yet, we do think that this problem deserves further study and that it can be studied using valid methods. On the basis of what we learned about the studies included in our review and based on pathophysiological concepts, we propose an optimal study design to answer the still open questions. We provide a guide for the design and execution of such a future study in the Appendix.

Conclusions

It is unclear to what extend subsequent neurodevelopment is impaired by neonatal hypoglycemia in the first week of life. Recommendations for clinical practice cannot be based on evidence because of a lack of valid empirical research. We propose a detailed design for a future methodological sound study to answer the still open question about the long-term prognosis of neonatal hypoglycemia. Our current proposal is open for debate, and we invite content experts and clinicians from all over the world to refine this design and to participate in a collaborative prospective meta-analysis.

Acknowledgment

This study was funded by the Centre for Clinical Practice Guidelines in the Academic Medical Centre (Amsterdam, Netherlands).

Appendix Optimal study design to establish the relationship between neonatal hypoglycemia and subsequent neurodevelopment

Decisions have to be made on (1) research questions, (2) design, (3) patient groups, (4) neonatal measurements, (5) glucose-assay method, (6) treatment, (7) outcome measurement instruments, (8) other risk factors for neurodevelopmental impairment to be measured, and (9) sample size. Table 4 summarizes the ingredients for an optimal study in this field.

Research Questions

1 What is the effect of various different glucose concentrations in the first three postnatal days on long-term neurodevelopment?
2 What is the effect of treatment with additional carbohydrates in neonates with moderate hypoglycemia on long-term neurodevelopment compared with expectant observation?

Design

The design should be that of a prospective cohort study to answer question 1, with a nested randomized, controlled clinical trial to answer question 2.

Patient (Sub)groups

High-risk patient groups for hypoglycemia with subsequent brain damage have not been defined in the available studies. Therefore, definition of high risk groups can currently be based on pathophysiological reasoning only. Hypoglycemia can cause brain damage, because glucose is the major fuel for brain cells. Alternative fuels for the perinatal brain are lactate and ketone bodies.[63;64] The immature brain is capable of using other organic metabolites like free fatty acids (FFA) and amino acids.[63;64] High risk groups for developing hypoglycemia are those infants with diminished glucose-production capacity (i.e., with impaired glycogenolysis or gluconeogenesis) and those with increased glucose utilization. Using this approach, infants who cannot rely on sufficient alternative fuel supply are at particular risk for subsequent brain damage. Thus, the most common risk groups for hypoglycemia with subsequent brain damage are term infants (SGA, LGA), and infants born preterm (SGA, AGA, LGA), and infants from diabetic mothers. These children are the prime target of future studies. Infants with symptomatic hypoglycemia must be analyzed separately. To attribute signs and symptoms to hypoglycemia, Whipples´s Triad must be fulfilled (i.e. a reliable low-blood glucose measurement, signs and symptoms consistent with hypoglycemia, and resolution of signs and symptoms after restoring blood glucose to normal values).[65]

Table 4 Optimal study design for elicitation of the neurodevelopmental outcome after episodes of neonatal hypoglycemia.

Research Questions	1 What is the effect of various different glucose concentrations in the first 3 postnatal days on long-term neurodevelopment? 2 What is the effect of treatment with additional carbohydrates in neonates with moderate hypoglycemia on long-term neurodevelopment compared to expectant observation?
Design	Prospective cohort study with a nested randomized, controlled clinical trial
Eligible patients	Asymptomatic neonates • Term infants, both LGA and SGA • Infants of diabetic mothers • Preterm infants, divided in subgroups according to gestational age, with a nonhypoglycemic control group matched for gestation and birth weight Symptomatic neonates
Measurements	Glucose, lactate and ketone bodies every 3 h, before feeding for 72 h and 1 h after treatment for hypoglycemia
Assay method	Plasma blood glucose by oxidase or hexokinase method, analyzed immediately or deproteinized and chilled
Treatment	Standardized treatment in 3 groups: A Very low glucose concentration (e.g. <1.8 mmol/L and/or symptomatic hypoglycemia); treatment always necessary B Moderate hypoglycemia, (e.g. 1.8-2.6 mmol/L; randomization in an intervention and a nonintervention group C No hypoglycemia, (e.g. >2.6 mmol/L); no treatment necessary Standardized treatment by raising the carbohydrate intake with 2 mg/kg per min
Other risk factors for impaired neurodevelopment	Cerebral hemorrhage, asphyxia, sepsis, respiratory distress syndrome, chronic lung disease, time spent on the ventilator, other neurological diseases
Outcome measurement instrument	• Validated neurological examination (e.g. Touwen) at 1,2, and 3 y • Validated neurodevelopmental test (e.g. Bayley Scales of Infant Development) at 1,2 and 3 y • Validated motor function test (e.g. Movement Assessment Battery for Children) • Validated Intelligence test (e.g. Wechsler Intelligence Scale for Children) at 7 y Outcome measurement should be blinded
Sample size	In order to exclude an IQ of >5 points lower in the hypoglycemic group and to anticipate on loss to follow-up, ideally 300 patients in each (sub)group should be included.

Measurements

Ideally, continuous measurements of glucose, alternative fuels (main: ketone bodies, lactate; additional: free fatty acids, amino acids), glucoregulatory hormones (insulin, glucagon, cortisol, growth hormone, catecholamines) and gluconeogenic precursors (glycerol, amino acids, lactate) should be collected. This approach will also reveal information on the pathogenesis of brain damage due to hypoglycemia and on the pathophysiology of hypoglycemia itself and could be helpful in identifying subgroups of

infants with a higher risk for hypoglycemia-induced brain damage. Because it will not be possible to measure all these variables in large groups of infants on a regular basis and, more important, these measurements are not available on short notice in daily clinical practice, the decision to treat hypoglycemia as yet is based on the blood glucose concentration only. Therefore, we propose the measurement of glucose, and the main alternative fuels ketone bodies and lactate every 3 hours before feeding for the first 72 hours of life. Number of episodes and depth and duration of hypoglycemia should all be recorded.

Assay Method

Faultless blood-sampling and handling techniques are critical, because small mistakes can cause major errors in the results. Therefore, the protocol should contain detailed instructions on blood-sampling, sample-handling, and analysis methods.

Treatment

Treatment for hypoglycemia should be standardized. Routine blood glucose measurements can result in: (1) very low glucose concentrations (e.g. <1.8 mmol/L), (2) moderately low glucose concentrations (e.g. 1.8-2.6 mmol/L), or (3) glucose concentrations considered 'safe' (e.g. >2.6 mmol/L). With very low glucose concentrations, and also in symptomatic hypoglycemia, treatment in the form of carbohydrate supplementation will always be instituted, and with glucose concentrations considered safe, no such treatment is necessary. Moderately low glucose concentrations are in the 'grey zone', and here the controversy about the necessity to treat these infants is maximal. Therefore, with moderately low glucose concentrations, the allocation to a carbohydrate-supplementation group and a non-intervention/close monitoring group could best be randomized. The short-term aim of the intervention is to keep the plasma glucose level above a very low glucose level for all children. An example of a standardized treatment intervention could be to raise the carbohydrate intake with a predefined amount of glucose (expressed in mg/kg per minute).

Other risk factors for neurodevelopmental impairment

Cerebral hemorrhage, asphyxia, sepsis, respiratory distress syndrome, chronic lung disease, time spent on the ventilator, other neurological diseases, etc, should be recorded prospectively according to accepted definitions and be explored as both confounders and effect modifiers of the central determinant-outcome relationship, preferably using multivariate-regression techniques.

Outcome Measurement Instruments

Animal studies have shown that severe hypoglycemia causes neuronal loss in the superficial cerebral cortex, the dentate gyrus, the hippocampus and the caudate nucleus.

The brainstem and the posterior fossa structures are apparently spared.[66] Therefore, we may expect severe hypoglycemia to be associated with both intellectual and motor deficits in survivors. Neurological function, neurodevelopment, and motor development should therefore be assessed at predefined ages using standardized and validated tests. Several studies have shown that cognitive delay, particularly in the mildest forms, is difficult to detect during early infancy. Therefore a validated Intelligence test should be performed at a suitable age. In any case, outcome assessors have to be blinded to the perinatal data.

Sample Size

For the intervention part of this prospective study (i.e. the comparison of more or less invasive treatment of children with moderately low glucose concentrations), the sample size should be based on the exclusion of a clinically important difference in neurodevelopmental outcome for the treatment groups. For example, for Bayley scales and Intelligence tests a difference of 1 SD (15 IQ points) is generally considered to be clinically important. Yet, because a 15-IQ point difference is rather large, we suggest to not tolerate an average IQ >5 IQ points lower in the nonintervention/close-monitoring group compared to the intravenous carbohydrate supplementation group. This 1-sided question could be answered in a noninferiority trial design. With 80% power and 1-tailed alpha of 0.05, 112 patients would need to be included in each group. To anticipate a loss to follow-up of, for example, 30%, at least 150 patients in each group should be included, which means 300 patients in total for each subgroup.

Reference List

1 Aynsley-Green A. Glucose, the brain and the paediatric endocrinologist. Hormone Research Vol 46 (1) (pp 8-25), 1996.

2 Cornblath M. Neonatal hypoglycemia 30 years later: does it injure the brain? Historical summary and present challenges. Acta Paediatr Jpn 1997; 39 Suppl 1:S7-11.

3 Cornblath M, Hawdon JM, Williams AF, Aynsley-Green A, Ward-Platt MP, Schwartz R et al. Controversies regarding definition of neonatal hypoglycemia: suggested operational thresholds. Pediatrics 2000; 105(5):1141-1145.

4 Cornblath M, Schwartz R, Aynsley-Green A, Lloyd JK. Hypoglycemia in infancy: the need for a rational definition. A Ciba Foundation discussion meeting. Pediatrics 1990; 85(5):834-837.

5 Pildes RS, Pyati SP. Hypoglycemia and hyperglycemia in tiny infants. Clinics in Perinatology Vol 13(2) (pp 351-375), 1986.

6 Laupacis A, Wells G, Richardson WS, Tugwell P. Users' guides to the medical literature. V. How to use an article about prognosis. Evidence-Based Medicine Working Group. JAMA 1994; 272(3):234-237.

7 Altman DG. Systematic reviews of evaluations of prognostic variables. BMJ 2001; 323(7306):224-228.

8 Abel HT, Lamme W, Theune L, Funfhausen H. [Hypoglycemia in newborn infants – a follow-up of twins]. Padiatr Grenzgeb 1987; 26(3):173-177.

9 Anderson JM, Milner RD, Strich SJ. Pathological changes in the nervous system in severe neonatal hypoglycaemia. Lancet 1966; 2(7459):372-375.

10) Anderson JM, Milner RD, Strich SJ. Effects of neonatal hypoglycaemia on the nervous system: a pathological study. J Neurol Neurosurg Psychiatry 1967; 30(4):295-310.

11 Barkovich AJ, Ali FA, Rowley HA, Bass N. Imaging patterns of neonatal hypoglycemia. AJNR Am J Neuroradiol 1998; 19(3):523-528.

12 Cakmakci H, Usal C, Karabay N, Kovanlikaya A. Transient neonatal hypoglycemia: cranial US and MRI findings. Eur Radiol 2001; 11(12):2585-2588.

13 Chiu NT, Huang CC, Chang YC, Lin CH, Yao WJ, Yu CY. Technetium-99m-HMPAO brain SPECT in neonates with hypoglycemic encephalopathy. J Nucl Med 1998; 39(10):1711-1713.

14 Cotte MF, Badinand P, Bethenod M. [Neonatal hypoglycemia: 16 cases]. Pediatrie 1969; 24(6):687-697.

15 Cowett RM, Howard GM, Johnson J, Vohr B. Brain stem auditory-evoked response in relation to neonatal glucose metabolism. Biol Neonate 1997; 71(1):31-36.

16 Creery RD. Hypoglycaemia in the newborn: diagnosis, treatment and prognosis. Dev Med Child Neurol 1966; 8(6):746-754.

17 Cresto JC, Abdenur JP, Bergada I, Martino R. Long-term follow up of persistent hyper-

insulinaemic hypoglycaemia of infancy. Arch Dis Child 1998; 79(5):440-444.

18 Dacou-Voutetakis C, Psychou F, Maniati-Christidis M. Persistent hyperinsulinemic hypoglycemia of infancy: Long term results. Journal of Pediatric Endocrinology & Metabolism Vol 11(SUPPL 1) (pp 131-141), 1998.

19 Dalgic N, Ergenekon E, Soysal S, Koc E, Atalay Y, Gucuyener K. Transient neonatal hypoglycemia—long-term effects on neurodevelopmental outcome. J Pediatr Endocrinol Metab 2002; 15(3):319-324.

20 Duvanel CB, Fawer CL, Cotting J, Hohlfeld P, Matthieu JM. Long-term effects of neonatal hypoglycemia on brain growth and psychomotor development in small-for-gestational-age preterm infants. J Pediatr 1999; 134(4):492-498.

21 Eeg-Olofsson O, Gentz J, Jodal U, Nilsson LR, Zetterstrom R. Neonatal symptomatic hypoglycemia. Some results from a follow-up study. Acta Paediatr Scand 1967;Suppl-6.

22 Fluge G. Neurological findings at follow-up in neonatal hypoglycaemia. Acta Paediatr Scand 1975; 64(4):629-634.

23 Francois R, Picaud JJ, Ruitton-Ugliengo A, David L, Cartal MJ, Bauer D. The newborn of diabetic mothers. Observations on 154 cases, 1958-1972. Biol Neonate 1974; 24(1):1-31.

24 Gentz J, Persson B, Zetterstrom R. On the diagnosis of symptomatic neonatal hypoglycemia. Acta Paediatr Scand 1969; 58(5):449-459.

25 Griffiths AD, Bryant GM. Assessment of effects of neonatal hypoglycaemia. A study of 41 cases with matched controls. Arch Dis Child 1971; 46(250):819-827.

26 Griffiths AD, Laurence KM. The effect of hypoxia and hypoglycaemia on the brain of the newborn human infant. Dev Med Child Neurol 1974; 16(3):308-319.

27 Haworth JC, McRae KN. Neonatal hypoglycemia: a six-year experience. J Lancet 1967; 87(2):41-45.

28 Haworth JC, McRae KN, Dilling LA. Prognosis of infants of diabetic mothers in relation to neonatal hypoglycaemia. Dev Med Child Neurol 1976; 18(4):471-479.

29 Hide DW. Symptomatic neonatal hypoglycaemia. Br J Clin Pract 1966; 20(10):533-535.

30 Hirabayashi S, Kitahara T, Hishida T. Computed tomography in perinatal hypoxic and hypoglycemic encephalopathy with emphasis on follow-up studies. J Comput Assist Tomogr 1980; 4(4):451-456.

31 Hunter A. Perinatal events and permanent neurological sequelae. N Z Med J 1968; 68(435):108-113.

32 Kinnala A, Rikalainen H, Lapinleimu H, Parkkola R, Kormano M, Kero P. Cerebral magnetic resonance imaging and ultrasonography findings after neonatal hypoglycemia. Pediatrics 1999; 103(4 Pt 1):724-729.

33 Knobloch H, Sotos JF, Sherard ES, Jr., Hodson WA, Wehe RA. Prognostic and etiologic factors in hypoglycemia. J Pediatr 1967; 70(6):876-884.

34 Koelz HR, Girard J. [On the syndrome of childhood-hypoglycemia. I. Definition of hypoglycemia in different age-groups and problems of neonatal hypoglycemia (author's transl)]. Monatsschr Kinderheilkd 1977; 125(7):694-701.

35 Kogut MD, Blaskovics M, Donnell GN. Idiopathic hypoglycemia: a study of twenty-six children. J Pediatr 1969; 74(6):853-871.

36 Koh TH, Aynsley-Green A, Tarbit M, Eyre JA. Neural dysfunction during hypoglycaemia. Arch Dis Child 1988; 63(11):1353-1358.

37 Koivisto M, Blanco-Sequeiros M, Krause U. Neonatal symptomatic and asymptomatic hypoglycaemia: a follow-up study of 151 children. Dev Med Child Neurol 1972; 14(5): 603-614.

38 Kumari S, Bhargava SK, Ahmed SH, Ghosh S. Transient symptomatic hypoglycemia in the newborn. Indian Pediatr 1971; 8(11):762-769.

39 Laurence KM, Griffiths AD. Proceedings: Hypoglycaemia in the newborn. Arch Dis Child 1974; 49(1):76-77.

40 Lucas A, Morley R, Cole TJ. Adverse neurodevelopmental outcome of moderate neonatal hypoglycaemia. BMJ 1988; 297(6659):1304-1308.

41 Menni F, de Lonlay P, Sevin C, Touati G, Peigne C, Barbier V et al. Neurologic outcomes of 90 neonates and infants with persistent hyperinsulinemic hypoglycemia. Pediatrics 2001; 107(3):476-479.

42 Murakami Y, Yamashita Y, Matsuishi T, Utsunomiya H, Okudera T, Hashimoto T. Cranial MRI of neurologically impaired children suffering from neonatal hypoglycaemia. Pediatr Radiol 1999; 29(1):23-27.

43 Nelson CA, Wewerka S, Thomas KM, Tribby-Walbridge S, deRegnier R, Georgieff M. Neurocognitive sequelae of infants of diabetic mothers. Behav Neurosci 2000; 114(5): 950-956.

44 Nunes ML, Penela MM, da Costa JC. Differences in the dynamics of frontal sharp transients in normal and hypoglycemic newborns. Clin Neurophysiol 2000; 111(2):305-310.

45 Persson B, Gentz J. Follow-up of children of insulin-dependent and gestational diabetic mothers. Neuropsychological outcome. Acta Paediatr Scand 1984; 73(3):349-358.

46 Pildes RS, Cornblath M, Warren I, Page-El E, Di Menza S, Merritt DM et al. A prospective controlled study of neonatal hypoglycemia. Pediatrics 1974; 54(1):5-14.

47 Pop L. [Brain damage in newborn infants with spontaneous hypoglycemia]. Monatsschr Kinderheilkd 1973; 121(7):317-319.

48 Priestley BL. Neurological assessment of infants of diabetic mothers in the first week of life. Pediatrics 1972; 50(4):578-583.

49 Shen EY. Neurosonographic findings in high risk neonates with hypoglycemia. Zhonghua Min Guo Xiao Er Ke Yi Xue Hui Za Zhi 1996; 37(4):248-252.

50 Singh M, Singhal PK, Paul VK, Deorari AK, Sundaram KR, Ghorpade MD et al. Neurodevelopmental outcome of asymptomatic & symptomatic babies with neonatal hypoglycaemia. Indian J Med Res 1991; 94:6-10.

51 Stenninger E, Eriksson E, Stigfur A, Schollin J, Aman J. Monitoring of early postnatal glucose homeostasis and cerebral function in newborn infants of diabetic mothers. A pilot study. Early Hum Dev 2001; 62(1):23-32.

52 Yamaguchi K, Mishina J, Mitsuishi C, Takamura T, Nishida H. Follow-up study of neonatal hypoglycemia. Acta Paediatr Jpn 1997; 39 Suppl 1:S51-S53.

53 Brand PL, Molenaar NL, Kaaijk C, Wierenga WS. Neurodevelopmental outcome of hypoglycaemia in healthy, large for gestational age, term newborns. Arch Dis Child 2005; 90(1):78-81.

54 Harris SR, Langkamp DL. Predictive value of the Bayley mental scale in the early detection of cognitive delays in high-risk infants. J Perinatol 1994; 14(4):275-279.

55 Moher D, Pham B, Klassen TP, Schulz KF, Berlin JA, Jadad AR et al. What contributions do languages other than English make on the results of meta-analyses? J Clin Epidemiol 2000; 53(9):964-972.

56 Pham B, Klassen TP, Lawson ML, Moher D. Language of publication restrictions in systematic reviews gave different results depending on whether the intervention was conventional or complementary. J Clin Epidemiol 2005; 58(8):769-776.

57 Cornblath M, Schwartz R. Outcome of neonatal hypoglycaemia. Complete data are needed. BMJ 1999; 318(7177):194-195.

58 Egger M, Smith GD, Phillips AN. Meta-analysis: principles and procedures. BMJ 1997; 315(7121):1533-1537.

59 Aynsley-Green A, Hawdon JM. Hypoglycemia in the neonate: current controversies. Acta Paediatr Jpn 1997; 39 Suppl 1:S12-S16.

60 Brand PL. What is the normal range of blood glucose concentration in healthy term newborns? Arch Dis Child Fetal Neonatal Ed 2004; 89(4):F375.

61 Hawdon JM. Hypoglycaemia and the neonatal brain. Eur J Pediatr 1999; 158 Suppl 1:S9-S12.

62 J.E.McGowan. Neonatal Hypoglycemia: Fifty Years Later, The Questions Remain the Same. NeoReviews 2004; 5(9):e363-e364.

63 Vannucci RC, Vannucci SJ. Glucose metabolism in the developing brain. Semin Perinatol 2000; 24(2):107-115.

64 Vannucci RC, Vannucci SJ. Hypoglycemic brain injury. Semin Neonatol 2001; 6(2):147-155.

65 Cornblath M, Ichord R. Hypoglycemia in the neonate. Semin Perinatol 2000; 24(2):136-149.

66 Auer RN, Siesjo BK. Hypoglycaemia: brain neurochemistry and neuropathology. Baillieres Clin Endocrinol Metab 1993; 7(3):611-625.

CHAPTER 4

State of the evidence on acute asthma management in children: a critical appraisal of systematic reviews

Nicole Boluyt, Johanna H. van der Lee, Virginia A. Moyer, Paul L.P. Brand, Martin Offringa

Submitted for publication

ABSTRACT

Objective To evaluate clinical, methodological and reporting aspects of systematic reviews on the management of acute asthma in children.

Methods We undertook a systematic review of systematic reviews on acute asthma management in children. We identified eligible reviews by searching the Cochrane Database of Systematic Reviews, the Database of Abstracts of Reviews of Effects, MEDLINE and EMBASE through March 2006. Data were extracted on clinical issues, methodological characteristics and results of the reviews. Methodological quality was assessed with the Overview Quality Assessment Questionnaire (OQAQ) and with additional questions on heterogeneity. Separate reporting on children in mixed adult-pediatric population reviews was assessed. Methodological quality of systematic reviews published in peer reviewed journals was compared to Cochrane reviews.

Results A total of 23 systematic reviews were included, 14 published in the Cochrane Library and 9 in peer reviewed journals. Eight reviews included children only, and 15 were mixed population reviews. The majority of reviews defined the study population as having 'acute asthma' without a more precise definition, and sixteen different health outcomes were reported. The overall quality according to the OQAQ was good, Cochrane reviews showing minimal flaws and journal reviews minor flaws (median scores 7 vs. 5; $P<0.001$). Results on children were reported separately in 8 of 15 mixed population reviews. Clinical heterogeneity was explored in only 2 of 23 reviews, and the methods used to identify and address heterogeneity were diverse.

Conclusions The methodological quality of both Cochrane and journal reviews on the management of acute asthma in children seems good, with Cochrane reviews being more rigorous. Yet, their usefulness for clinical practice is hampered by a lack of clear definitions of included populations, clinical important health outcomes, and separate reporting on children in mixed reviews. A major threat to these reviews' validity is the insufficient identification and handling of heterogeneity.

Introduction

Acute asthma is a common reason for coming to emergency departments (ED). In the United States alone, acute asthma accounts for almost 2 million ED visits per year.[1] Approximately 10-20% of these patients require hospital admission, and another 10-20% will relapse within the subsequent two weeks.[2;3] These outcomes depend on the treatment prescribed in the ED and at discharge, which consists mainly of inhaled beta-2 agonists and systemic corticosteroids. A variety of other agents for acute treatment can be used, including anticholinergics, magnesium sulphate, heliox and theophylline.

During the past two decades, systematic reviews have gained popularity as a way of coping with increasing amounts of information about new devices and drugs. Systematic reviews synthesize large amounts of research evidence to help bridge the gap between evidence from single studies and clinical practice. In addition, they guide future research into the information gaps that are identified through the review process. Systematic reviews may include meta-analysis, i.e. the statistical combination of results of several independent studies to produce a single estimate of the effect of a particular healthcare intervention.[4;5] Properly performed, systematic reviews represent the highest level of evidence available to clinicians.

If clinicians are to have confidence that the results of systematic reviews can be used to guide clinical practice and the research agenda, then these reviews need to be of high quality. Unfortunately, most systematic reviews and meta-analyses published in peer reviewed journals have been shown to have methodological deficiencies that limit their validity and the applicability of their results in practice.[6-12] So far, these alarming results all come from studies on health care interventions in *adult* patients.

In pediatrics, two additional challenges face the clinician who wishes to use systematic reviews as a tool for guiding clinical decision making. First, and in contrast to systematic reviews of studies in adults, the volume of trials and the number of children included is often small making definite conclusions difficult. Second, many reviews include both adults and children (mixed populations reviews). Given the differences in response to treatments between adults and children, pooling of data from mixed populations could lead to the incorporation of invalid findings for pediatric clinical practice. On the other hand, if mixed reviews are disregarded altogether important information could be missed.[13]

We undertook a comprehensive search for systematic reviews on pediatric acute asthma management issues with the following goals:

1 To provide an overview of the currently available evidence syntheses on acute asthma treatment in children; examining the subject content, clinical aspects, methodological aspects, and the results of these systematic reviews
2 To evaluate the quality of these systematic reviews, compare the methodological quality of Cochrane reviews with reviews published in peer reviewed journals, and

to assess how data on children with acute asthma are reported in mixed reviews
3 To identify areas for improvement in the methods and reporting of these reviews

Methods

Literature search and study selection

We identified eligible studies by searching the Cochrane Database of Systematic Reviews, the Database of Abstracts of Reviews of Effects (DARE) (Cochrane Library; issue 1, 2006), MEDLINE and EMBASE through March 2006, using the search term 'asthma', limits 'All Child: 0-18 years'. In MEDLINE and EMBASE we added a comprehensive search strategy for systematic reviews and meta-analysis. In addition, the National Guideline Clearinghouse (www.guideline.gov) was searched for evidence-based practice guidelines ('acute asthma' AND 'child') and reference lists of the guidelines were checked for systematic reviews. Abstracts of potentially eligible reviews were assessed independently by two reviewers for inclusion. The final decision on inclusion in the review was made by consensus.

Inclusion criteria

To be included, a report had to explicitly indicate the use of a strategy for locating evidence or had to be described as a systematic review or a meta-analysis on the treatment of acute asthma in children, which we defined as individuals $\leq$18 years of age. We also included systematic reviews and meta-analyses on mixed populations.

Data extraction

Two reviewers extracted data independently by using a data collection form specifically designed for this study and piloted on five systematic reviews not included in this sample. Data were extracted on general characteristics of the reviews (publication year, source, source of funding); clinical issues (population, definition of acute asthma, intervention, and outcomes reviewed); methodological characteristics (design, methodological quality and funding of the included trials, heterogeneity testing); and results. Differences were resolved by consensus.

Methodological quality was assessed with the Overview Quality Assessment Questionnaire (OQAQ), the only validated instrument available for the critical appraisal of review articles.[14;15] The OQAQ checklist addresses 5 different aspects of scientific quality of a systematic review (search strategy, selection, quality assessment, pooling and results) and evaluates the overall scientific quality of the review on a 7-point scale, with 1 and 2 indicating extensive flaws, 3 and 4 indicating major flaws, 5 and 6 indicating minor flaws, and 7 indicating minimal flaws. To further assess methodological quality, we added items that are not fully covered by the OQAQ on the presence of hetero-

geneity: 1) Was the presence or absence of heterogeneity identified?; 2) Was clinical heterogeneity explored?; 3) Were *a priori* subgroup analyses by age and/or asthma severity planned?; 4) Were the reasons for statistical heterogeneity, if present, explored?; 5) Was heterogeneity, if present, correctly accounted for either by using alternative data-pooling techniques or by refraining from pooling?

Analysis

The main analysis of the data was descriptive. To test the difference in OQAQ total scores between Cochrane and journal systematic reviews, a Mann-Whitney test was performed, since data were not normally distributed.

Results

Search and general characteristics

We identified 29 reviews on pediatric acute asthma management; 4 were published both in a peer reviewed journal and in the Cochrane Library;[16-19] one included only adult patients[20] and one appeared to be a narrative overview instead of a systematic review.[21] Twenty-three citations met our inclusion criteria,[22-44] the earliest published in 1992.[39] Fourteen were Cochrane reviews[24;25;27;29;30;32;34;36;37;40-44] and 9 were published in peer reviewed journals ('journal reviews').[22;23;26;28;31;33;35;38;39]

Patients, interventions and outcomes

Table 1 shows clinical characteristics and main results of the reviews, by intervention.

Eight reviews included children only, and 15 were mixed population reviews. The number of pediatric studies included in the reviews varied from 1 to 19 and the total number of children included varied from 11 to 1712. In all reviews the study population was defined as having 'acute asthma'. Four reviews (17%) described the definition of acute asthma and severity; it was either based on a clinical score,[25] on a combination of a clinical score and FEV_1[36] or on FEV_1 or PEF.[24;38] Nineteen reviews (83%) did not report a definition of acute asthma or the inclusion criteria for eligible patients.

Sixteen different outcomes were reported; 22 reviews reported pulmonary function tests; 18 hospitalization rate; 12 adverse effects; 10 asthma symptoms; 7 length of hospitalization; 5 vital signs; 4 ED treatment duration; 4 relapse to additional care; 4 β2-agonist use; 4 heart rate; 4 mechanical ventilation; 3 costs; 2 corticosteroid use; 2 quality of life; 2 ICU admission, and 1 arterial blood gas values.

Methodological characteristics

Tables 2, 3 and 4 summarize the component scores of the OQAQ and the other methodological quality items for reviews published in the Cochrane Library and in other journals.

Table 1 Clinical characteristics of included systematic reviews on acute asthma management, by intervention

First Author (source, year updated)	Population	Clinical Characteristics	Intervention	Comparison	Outcomes*	Results (95% CI)	OQAQ score (0-7)
BETA2-AGONISTS							
Travers[44] (Cochrane 2001)	Adults and children 0.6-18 y	Severe acute asthma presenting to ED	Iv selective or non-selective β2-agonists	Inhaled β2-agonists, placebo, other iv bronchodilators (methylxanthines)	E, F, L, M,	E PEFR: WMD -24.7 l/min (2.9, -52.3) (mixed) F autonomic side effects: OR 2.2 (0.9-5.7) (mixed) L: 4.5 beats/minute (-4.9-14.0)	6
Camargo Jr[25] (Cochrane 2003)	Adults and children 2-18 y	Children presenting to ED with moderate-severe asthma exacerbation (clinical score ≥8)	Continuous inhaled β2-agonists (or nebulization ≥4 per h)	Intermittent inhaled β2-agonists	Primary: E Secondary: A, C, F, P	A: NNT 10 (6-34) (mixed) A: RR 0.68 (0.5-0.9) (mixed) E % pred FEV1 : WMD 0.28 (0.03-0.5) (mixed) F: WMD 0.2 (0.03-1.6) F tremor: RR 0.81 (0.5-1.3)	7
COMBINATION BETA2-AGONISTS AND ANTICHOLINERGICS							
Osmond[35] (1995)	Children only 1-17 y	Acute, unprovoked attack of asthma	Inhaled ipratropium-bromide and a β2-agonist	Placebo and a β2-agonist	A, B, E, K	E %pred FEV1: WMD 12.5% (6.6-18.4)	5
Plotnick[36] (Cochrane 2000)	Children only 13 mo-17 y	Acute unprovoked asthma exacerbation presenting to ED Mild: clinical score 1-3 Moderate: FEV1 50-70% or clinical score 4-6 Severe: FEV1 <50% or clinical score 7-9	Repeated doses of nebulized or inhaled short-acting anti-cholinergic and β2-agonists	Only short acting β2-agonists	Primary: A Secondary: E, F, G	A: NNT 12 (8-32) A: NNT 7 (5-20) severe asthma E %pred FEV1: WMD 9.68 % (5.7-13.7)	7
Aaron[22] (2001)	Adults and children 2-18 y	Acute airflow obstruction in pediatric asthma	Ipratropium + β2 agonist	β2-agonist	A, C, E	A: NNT 6.6 (3.7-29.4)	2
Rodrigo[38] (2005)	Adults and children 18 mo-17 y	Acute asthma exacerbation Moderate: FEV1 or PEF 50-70% Severe: FEV1 or PEF <50%	Inhaled β2-agonist and anticholinergics (>2 doses)	Inhaled β2-agonist	A, E	A: NNT 13 (9-28) A: NNT 7 (4-16) severe asthma A: RR 0.73 (0.63-0.85) E change in FEV1: WMD 16.3% (8.2-24.5)	7

First Author (source, year updated)	Population	Clinical Characteristics	Intervention	Comparison	Outcomes*	Results (95% CI)	OQAQ score (0-7)
INHALATION DEVICES							
Amirav[23] (1997)	Children only 0.5-18 y	Children with acute asthma	Metered dose inhalers	Smart volume nebulizers or no comparison	A, E, O, P	No pooled effect estimate	2
Cates[27] (Cochrane 2006)	Adults and children >2 y	Children with acute asthma. Presenting in the community or hospital emergency setting, or children already admitted to hospital	Any β-agonist given by MDI with any holding chamber	Any nebulizer	Primary: A, B Secondary: C, E, L	A: RR 0.65 (0.4-1.06) C: WMD -0.47 h (-0.58- -0.37) L: WMD -7.59% (-9.94- -5.24%)	7
Castro-Rodriguez[26] (2004)	Children only <5 y	Children with acute exacerbation of wheezing or asthma	Any β-agonist given by MDI with valved holding chamber	Any nebulizer	A	A: NNT 10 (6-26) A: OR 0.42 (0.24-0.72) A: OR 0.27 (0.13-0.54) moderate/ severe	7
STEROIDS							
Rowe[39] (1992)	Adults and children 2 y– not stated	Moderate to severe exacerbations	Glucocorticosteroids (IV, IM, oral)	Not stated	A, D, E, I, J	A: NNT 6-11 A: OR 0.07-0.42 D: OR 0.15 (0.04-0.44) after 7-10 d (mixed) E: ES -0.073 SD (-0.39-0.25) 1SD=18% pred FEV1 parenteral vs oral (mixed)	6
Edmonds[29] (Cochrane 2003)	Adults and children ≥2 y	Patients discharged from ED following assessment and treatment for acute asthma	ICS	Oral corticosteroid	A, D, E, F, H, I, J	D: OR 1.0 (0.66-1.52) after 7-10 d (mixed) E absolute PEFR: 11.0 L/min (-1-23) J: SMD -0.1 (-0.4-0.1)	7
Rowe[42] (Cochrane 2001)	Adults and children Age not stated	Patients presenting to ED (and discharged from ED) with acute exacerbation of asthma	Corticosteroids (IV, IM, oral)	Placebo	Primary: D Secondary: A, E, F, H, I	A: NNT 16 (7-125) (mixed) D: NNT 13 (7-91) (7-10 d) (mixed) D: OR 0.35 (0.17-0.73) (mixed) F: OR 0.94 (0.42-2.13) H: WMD -3.3 activations/d (-5.5- -1.0)	7

First Author (source, year updated)	Population	Clinical Characteristics	Intervention	Comparison	Outcomes*	Results (95% CI)	OQAQ score (0-7)
Rowe[41] (Cochrane 2002)	Adults and children ≥1 y	Patients with acute asthma presenting to ED	Corticosteroids (IV, IM, oral) early (<1 h) in the ED treatment	Placebo	Primary: A Secondary: E, F	A: NNT 8 (5-21) (mixed) A: OR 0.40 (0.17-0.94) A: OR 0.24 (0.11-0.53) oral steroids A: OR 0.68 (0.39-1.21) IV steroids (mixed)	7
Smith[43] (Cochrane 2003)	Children only 1-18 y	Severe acute asthma. Only patients treated in an ED and requiring hospital admission	Corticosteroids (IV, IM, oral)	Placebo or inhaled steroids	B, C, D, E, F, I	B: WMD -8.75 h (-19.23-1.74) C: NNT 3 (2-8) discharge after 4 h C: OR 7.00 (2.98-16.45) D: NNT 3 (2-7) D: OR 0.19 (0.07-0.55) E %pred PEFR: WMD 7.21 (-7.01-21.25)	7
Edmonds[30] (Cochrane 2003)	Adults and children >2 y	ED treatment for acute asthma	ICS (+ systemic corticosteroids)	Placebo or systemic corticosteroids	A, E, F, I, K	A: OR 0.14 (0.03-0.60) placebo A: OR 0.89 (0.2-4.5) systemic steroids	7
HELIOX							
Ho[33] (2003)	Adults and children 16 mo-18 y	Acute asthma requiring hospital treatment	Any mixture of helium and oxygen	Not stated	B, E, I, O	E %pred PEFR: WMD 3% (92% CI -2-8%) (mixed)	4
Rodrigo[37] (Cochrane 2003)	Adults and children 5-18 y	Asthma exacerbation presenting to ED or equivalent care setting	Inhaled heliox	Oxygen or air	Primary: E Secondary: A, F, I, L, O	E: SMD 0.13 (-0.09-0.34) (mixed) L: SMD 7.67 (0.79-14.55) (mixed)	7
MAGNESIUM SULFATE							
Rowe[40] (Cochrane 2000)	Adults and children Age not stated	Patients presenting to an ED for treatment of acute asthma	Intravenous magnesium sulfate	Placebo	Primary: A Secondary: E, F, K, L	A: OR 0.31 (0.09-1.02) (mixed) A: OR 0.10 (0.04-0.27) severe asthma (mixed) A: OR 1.36 (0.72-2.55) mild-moderate asthma (mixed) E %pred FEV1: WMD 9.8% (3.8-15.8) (mixed) L: WMD 5.6 beats/minute (-1.5-12.7) (mixed)	7
Blitz[24] (Cochrane 2005)	Adults and children 2-18 y	Acute severe asthma defined as FEV1 or PEF <50%	1 Inhaled magnesium sulfate plus β2-agonist or 2 Inhaled magnesium sulfate	β2-agonist	A, E, F, K	1 A: RR 2.00 (0.19-20.93) E: SMD 0.27 (-0.12-0.66) (mixed) E: SMD 0.55 (0.12-0.98) severe (mixed) F: RD 0.00 (-0.03-0.03) (mixed) 2 A: RR 0.50 (0.04-6.12) (mixed)	7

First Author (source, year updated)	Population	Clinical Characteristics	Intervention	Comparison	Outcomes*	Results (95% CI)	OQAQ score (0-7)
Cheuk(28) (2005)	Children only <18 y	Acute moderate to severe asthmatic attacks, inadequate response to first line treatment	Intravenous magnesium sulfate	Placebo	A, E, I, N	A: NNT 4 (3-8) E PEFR: 8.58% (0.94-16.22) I: 1.33 (0.31-2.36)	5
OTHER INTERVENTIONS							
Goodman(31) (1996)	Children only <18 y	Children hospitalized with acute asthma	Theophylline or aminophylline	IV saline or dextrose solution, IV albuterol	B, E, H, I, O	B: WMD -0.31 d (-0.3-0.05) E: ES 1.6 SD (-2.6-5.9)	4
Graham(32) (Cochrane 2005)	Adults and children Age not stated	Patients presenting to an ED with acute asthma	Antibiotics (oral or intravenous)	Placebo	Primary: A, B Secondary: E, K, P	B: WMD -0.10 (-0.53-0.33) E %pred FEV1: WMD 9.8 (-3.58-23.18) at 24 h	6
Mitra(34) (Cochrane 2005)	Children only 2-17 y	Acute severe asthma or status asthmaticus attending emergency departments or hospital wards	IV aminophylline (in addition to maximised inhaled broncho-dilators and gluco-corticoids)	Placebo	Primary: E Secondary: B, E, F, H, I, N	E % pred FEV1: WMD 8.37% (0.82-15.92) at 6-8 h B: WMD - 2.1 h (-9.45-5.25) F vomiting: RR 3.59 (2.15-6.33)	7

* Outcomes reported in Methods Section of the review

Abbreviations: y: year; m: month; d: day; ED: Emergency Department; PEFR: peak expiratory flow rate; WMD: weighted mean difference; OR: odds ratio; NNT: number needed to treat; RR: relative risk; FEV1: forced expiratory volume at 1 second; MDI: metered-dose inhaler; ES: effect size; ICS: inhalation corticosteroids; SMD: standardized mean difference

Outcomes: A: hospitalization rate; B: length of hospitalization; C: Emergency Department treatment times; D: relapse to additional care; E: pulmonary function tests; F: adverse effects; G: corticosteroids use; H: β2-agonist use; I: asthma symptoms; J: quality of life; K: vital signs; L: heart rate; M: arterial gas values; N: ICU admission; O: mechanical ventilation; P: costs

Table 2 Methodological quality of included systematic reviews according to Overview Quality Assessment Questionnaire (OQAQ)

Items OQAQ*	1	2	3	4	5	6	7	8	9	10
Beta2-agonists										
Travers (Cochrane 2001)	+	+	+	+	+	+	+	+	+	7
Camargo Jr (Cochrane 2003)	+	+	+	+	+	+	+	+	+	7
Combination beta2-agonists and anticholinergics										
Osmond (1995)	+	-	+	+	+	-	+	+	+	5
Plotnick (Cochrane 2000)	+	+	+	+	+	+	+	+	+	7
Aaron (2001)	-	?	-	?	-	-	-	-	P	2
Rodrigo (2005)	+	+	+	+	+	+	+	+	+	7
Inhalation devices										
Amirav (1997)	+	-	P	?	-	?	-	-	P	2
Cates (Cochrane 2003)	+	+	+	+	+	+	+	+	+	7
Castro-Rodriguez (2004)	+	+	+	+	+	+	+	+	+	7
Steroids										
Rowe (1992)	+	-	+	+	+	-	+	?	+	5
Edmonds (Cochrane 2000)	+	+	+	+	+	+	+	+	+	7
Rowe (Cochrane 2000)	+	+	+	+	+	+	+	+	+	7
Rowe (Cochrane 2002)	+	+	+	+	+	+	+	+	+	7
Smith (Cochrane 2002)	+	+	+	+	+	+	+	+	+	7
Edmonds (Cochrane 2003)	+	+	+	+	+	+	+	+	+	7
Heliox										
Ho (2003)	+	+	+	+	+	?	P	?	+	4
Rodrigo (Cochrane 2003)	+	+	+	+	+	+	+	+	+	7
Magnesium sulfate										
Rowe (Cochrane 2000)	+	+	+	+	+	+	+	+	+	7
Blitz (Cochrane 2005)	+	+	+	+	+	+	+	+	+	7
Other interventions	+	+	+	+	+	+	+	+	+	7
Goodman (1996)	+	-	+	?	P	?	+	?	+	4
Graham (Cochrane 2001)	+	+	+	?	+	+	+	+	+	6
Mitra (Cochrane 2005)	+	+	+	+	+	+	+	+	+	7

* Items: 1. Search methods stated; 2. Search comprehensive; 3. Inclusion criteria reported; 4. Selection bias avoided; 5. Validity criteria reported; 6. Validity assessed appropriately; 7. Combining methods reported; 8. Findings combined appropriately; 9. Conclusions supported by data; 10. Overall quality
+: yes; -: no; p: partially; ?: can't tell

OQAQ

The median OQAQ score of all reviews was 7 (range 2-7), indicating minimal flaws. Cochrane reviews had a median score of 7 (range 6-7), indicating minimal flaws, and journal reviews a median score of 5 (range 2-7), indicating minor flaws (P<0.001). Seven out of 9 journal reviews exhibited minor to minimal flaws; two had extensive flaws.

Heterogeneity

The presence or absence of heterogeneity was mentioned in 20 reviews but only 2 reviews stated in their methods the intent to explore *clinical* heterogeneity between

Table 3 Overview Quality Assessment Questionnaire (OQAQ) component scores and other methodological considerations for Cochrane and journal reviews. Values are numbers of affirmative answers/number of reviews

Methodological Factor	All	Cochrane reviews	Peer reviewed journals
OQAQ Component Scores			
Search methods used to find evidence stated	22/23	14/14	8/9
Search for evidence reasonably comprehensive	18/23	14/14	4/9
Criteria used for deciding which studies to include reported	21/23	14/14	7/9
Bias in the selection of studies avoided	19/23	13/14	6/9
Criteria used for assessing the validity of the included studies reported	20/23	14/14	6/9
Validity of all studies referred to in the text assessed appropriately	17/23	14/14	3/9
Methods used to combine the findings of the relevant studies reported	20/23	14/14	6/9
Findings of relevant studies combined appropriately	18/23	14/14	4/9
Conclusions made by the author(s) supported by data or analysis	21/23	14/14	7/9
Overall quality (median score)	7	7	5*
Other Methodological Considerations			
Children only included	8/23	3/14	5/9
Children reported separately in mixed reviews	8/15	5/11	3/4
No potential conflict of interest	9/23	8/14**	1/9‡
Updating#	12/23	8/14	4/9

* Mann-Whitney test: P-value (2-sided) <0.001

** 6 stated that the authors received industry grants, but the development of the review was not industry sponsored

‡ 8 did not state potential conflict of interest

\# Updated <3 years ago or published <3 years ago

studies. Of the 20 reviews that considered heterogeneity, 13 pre-specified subgroup analyses or sensitivity analyses. Eight of 15 mixed reviews reported results on children separately. Ten of the 15 mixed reviews pre-specified a subgroup analysis for age and 5 did not. However, 6 out of 10 reviews that planned a subgroup analysis for age did not actually report children separately either because no heterogeneity was found, or the number of studies in children was considered insufficient. A subgroup analysis based on asthma severity was pre-specified in 11 of 23 reviews. Most reviews accounted for heterogeneity either by sensitivity or subgroup analyses (n=5), by using the more conservative random effects model (n=8), or by refraining from pooling (n=3).

Funding

None of the systematic reviews provided information about the sources of funding of the individual studies that it included. Seven of the 14 Cochrane reviews stated no potential conflict of interest; the 7 others stated that the authors had received grants from industry. Eight of nine journal reviews did not provide any information about a potential conflict of interest.

Table 4 Methodological characteristics of included systematic reviews on acute asthma management

Author (year)	Design included studies	Included pediatric studies (total studies included)	Children included (n=)	Funding	Assessment of allocation concealment and blinding	Pre-specified subgroup analysis	Heterogeneity investigating its cause by	Heterogeneity accounted for by
Travers (Cochrane 2001)	(quasi) RCT	3 (15)	-	Independent	Yes	Age and asthma severity	Subgroup analyses and sensitivity analyses	Sensitivity analyses
Camargo Jr (Cochrane 2003)	RCT	1 (8)	70	Independent	Yes	Age and asthma severity	Subgroup analyses and sensitivity analyses	Random effects model
Osmond (1995)	RCT	6	285	-	-	-	-	NA, not present
Plotnick (Cochrane 2000)	RCT	13	1712	Industry grants for authors	Yes	Asthma severity	Sensitivity analyses	Random effects models
Aaron (2001)	RCT (and SR)	2 (1 SR)	-	-	-	-	-	-
Rodrigo (2005)	RCT	16 (32)	1564	Industry grants for authors	Yes	-	Subgroup analyses and sensitivity analyses	NA, not present
Amirav (1997)	RCT	10	582	-	-	-	-	-
Cates (Cochrane 2003)	RCT	19 (27)	1260	Independent	Yes	Age	Sources of hetero-geneity identified	Random effects model, or not pooled
Castro-Rodriguez (2004)	RCT	6	491	-	Yes	Asthma severity	Sensitivity analysis	NA, not present
Rowe (1992)	(quasi) RCT	9 (30)	-	-	-	-	Subgroup analyses	Subgroupanalysis
Edmonds (Cochrane 2000)	(quasi) RCT	4 (10)	404	Industry grants for authors	Yes	Age and asthma severity	Sensitivity analyses	Random effects model
Rowe (Cochrane 2000)	RCT	2 (7)	-	Industry grants for authors	Yes	-	Subgroup analyses and sensitivity analyses	NA, not present
Rowe (Cochrane 2002)	(quasi) RCT	6 (12)	448	Industry grants for authors	Yes	Age and asthma severity	Subgroup analyses and sensitivity analyses	Subgroupanalyses and random effects model

	Design included studies	Included pediatric studies (total studies included)	Children included (n=)	Funding	Assessment of allocation concealment and blinding	Pre-specified subgroup analysis	Heterogeneity investigating its cause by	Heterogeneity accounted for by
Smith (Cochrane 2002)	RCT	7	426	Industry grants for authors	Yes	-	-	NA, not present
Edmonds (Cochrane 2003)	(quasi) RCT	3* (7)	-	Industry grants for authors	Yes	Age and asthma severity	Sensitivity analyses	Random effects model
Ho (2003)	All designs	2 RCTs	-	-	-	-	-	No pooling
Rodrigo (Cochrane 2003)	RCT	1 (6)	11	Independent	Yes	Age	NA (not present)	NA, not present
Rowe (Cochrane 2000)	(quasi) RCT	2 (7)	78	Industry grants for authors	Yes	Age and asthma severity	Subgroup analyses and sensitivity analyses	-
Blitz (Cochrane 2005)	(quasi) RCT	3 (6)	135	Independent	Yes	Age and asthma severity	Subgroup analyses and sensitivity analyses	Subgroupanalyses, sensitivity analyses and random effects model
Cheuk (2005)	CCT	5	182	-	Yes	Age and asthma severity	Exploring clinical heterogeneity and sensitivity analyses	Yes
Goodman (1996)	RCT	6	164	-	-	-	-	No pooling
Graham (Cochrane 2001)	RCT	1 (2)	37	Independent	Yes	-	-	-
Mitra (Cochrane 2005)	RCT	7	380	Independent	Yes	Asthma severity	Subgroup analyses and sensitivity analyses	Subgroupanalysis and random effects model

Abbreviations: - : no or not stated; NA: not applicable

* Some pediatric studies mentioned to be included in the methods were not reported in the results

Updating

Of the Cochrane reviews, 6 out of 14 were more than three years old (range 3-6 years). Publication dates of the journal reviews ranged from 1992-2005, with 5 out of 9 reviews being published more than three years prior to our search.

Discussion

Our assessment of the existing volume and quality of systematic reviews on the management of acute asthma in children presenting to the ED provides useful insights with regard to the current scientific basis for clinical decision making and the research agenda. In the following we will discuss our findings on the clinical and methodological characteristics of the reviews, and give recommendations for future trials and systematic reviews in this field.

Clinical characteristics of included systematic reviews

The majority (83%) of the reviews did not state a clear definition of acute asthma, but instead used whatever definition was used by the included primary studies. This holds two potential problems. As treatments may have different effects sizes in patients with different clinical presentation of acute asthma (e.g. severe asthma), pooling of these potentially clinical heterogenous studies results in an invalid overall effect size.[45] The second problem is that without specifically stating population characteristics, clinicians will not be able to evaluate external validity, i.e. the extent to which the results are generalizable and applicable to their own population of children with acute asthma.

Many different health outcomes were reported, the most frequently reported being pulmonary function tests, hospitalization rate and adverse effects. The reporting of pulmonary function tests in the individual trials was heterogeneous (e.g. % predicted or % change from baseline), and pooling was often not possible. Recently, new evidence has documented an additional threat to the validity of systematic reviews; selective reporting of trial outcomes within published studies.[46;47] The most common reason for nonpublication of selective outcomes is lack of statistical significance. As selective reporting is widespread and can change the conclusions of systematic reviews it should be routinely investigated in future systematic reviews.[47]

Methodological characteristics of included systematic reviews

The OQAQ scores of the included systematic reviews published in peer reviewed journals exceeded those reported in most other (adult) studies.[6-12] A critical evaluation of 50 systematic reviews and meta-analyses on asthma treatment in adults showed that 80% of the reviews published in peer reviewed journals had major flaws, compared to 23% in our sample.[11]

Several sources of bias that can influence the results of systematic reviews have been identified and studied. One of these is the presence of publication bias, known to occur in systematic reviews because smaller studies or studies yielding non-significant results are less likely to be published, leading to an overestimation of the treatment effect when only published studies are included.[48] For that reason a comprehensive search should be performed, including published and unpublished research (item 2 of the OQAQ (Table 2)). All Cochrane reviews in our sample searched the three major electronic databases (MEDLINE, EMBASE and The Cochrane Library), but the majority (56%) of the journal reviews only searched MEDLINE. A method to identify unpublished studies was stated by all Cochrane reviews and 33% of the journal reviews. The only satisfactory way to address publication bias and to improve the quality of the conduct, analysis and reporting of studies, meta-analyses and systematic reviews is through prospective registration of trials.[49] Unfortunately, these trial registers are still voluntary. The confidence in results generated by systematic reviews depends to a large extent on the quality of the primary studies included. Including studies of low quality in systematic reviews have been shown to exaggerate treatment effectiveness by 30-50%.[50] The two most important items that may lead to overestimation of treatment effect are a lack of allocation concealment and a lack of blinding of effect assessors.[50;51] Almost all reviews (91%) included randomized controlled trials, but 33% also included quasi-randomized controlled trials, which – by definition – lack allocation concealment. All Cochrane reviews and 33% of the journal reviews assessed allocation concealment and blinding in the individual studies (item 6 of the OQAQ).

Heterogeneity

Statistical pooling of study results can produce serious errors in the estimation of effect sizes, especially when combining heterogeneous results of individual studies.[49;52] Various types of heterogeneity can be distinguished; variability in the participants, interventions and outcomes in studies is described as *clinical heterogeneity*, and variability in trial design and quality is described as *methodological heterogeneity*. The pooling of data from heterogeneous primary studies can lead to the occurrence of *statistical heterogeneity*, which is variability in estimates of effect sizes across individual studies.[49] Meta-analysis should only be considered when a group of trials is sufficiently homogeneous in terms of participants, interventions, health outcomes, and methodology.

The majority of the reviews in our study (87%) identified the presence of heterogeneity, but only two reviews considered clinical heterogeneity. The statistical tests to detect heterogeneity that were used by 17 of 20 reviews that considered heterogeneity are known to have limited power when a small number of studies is included, which is often the case in pediatric systematic reviews.[52] Thus, the presence of heterogeneity is probably underestimated. The more recently developed measure I^2 is used to assess the impact of heterogeneity on a meta-analysis, independent of the number of studies.[53]

Three of the most recent reviews used this measure.[26;34;38] The majority (87%) of reviews pooled individual study results, despite possible heterogeneity in study population due to lack of clear definitions for acute asthma. Additional subgroup analyses, such as for asthma severity and age (for mixed reviews), should be specified a priori.[52] Only half of the reviews pre-specified subgroups based on these characteristics. Thus, and despite a good overall score on the OQAQ, we think that the current main threat to the validity of pediatric systematic reviews (both Cochrane and journal reviews) is how they deal with clinical and statistical heterogeneity.

Funding of primary trials and of reviews

Published trials sponsored by industry have been shown to exaggerate treatment effects and come to a positive conclusion in favor of the sponsor's drug five times more often than do not-for-profit-sponsored trials.[54] None of the systematic reviews in this field stated the source of funding of the individual studies. Like individual studies, reviews performed with industry sponsorship have been shown to have an increased risk of producing results in favour of the interventions promoted by the sponsors.[55;56] In the present series, all Cochrane reviews addressed potential conflict of interest. In contrast, 88% of the journal reviews did not give any information regarding a potential conflict of interest.

Updating of reviews

Reviews should be updated regularly as the emerging of new trials may change the conclusion of a systematic review. Cochrane reviews are supposed to be updated every two years.[49] In the Cochrane reviews identified, 43% had not been updated since 3 years. The reason for this could be that no new studies have been published. However, to inform the reader this should be stated in a bi-annually updated review. Hard copy journal reviews cannot be updated and should be appraised critically for their freshness; 56% were published more than 3 years ago (as per July 2006).

Mixed populations

Only 8 of the 15 reviews including both adults and children reported results for children separately. The reasons stated were either that there were not enough data on children or that 'no heterogeneity was found'. As pointed out earlier, in most cases the number of included studies was small in which case the power of the traditional statistical test to detect heterogeneity is low.

Mixed reviews present the problem of generalizability of study results: depending on the subject matter pediatricians may not be comfortable using mixed data results and applying them to the pediatric population. A recent study that investigated differences in effect sizes between adults and children in mixed populations systematic reviews could not exclude clinically important differences due to lack of power.[57] Yet,

there is evidence that in children with asthma the mechanisms of drug action may be different from adults.[58]

Cochrane reviews versus peer reviewed journal reviews

This study confirms that Cochrane reviews tend to be more rigorous in their methods and reporting than reviews published in peer reviewed journals.[11;59] The reason is probably that Cochrane reviews are developed following standardized instructions, have a peer reviewed protocol and receive input from different reviewers at different stages of development. Two systematic reviews published in peer reviewed journals with high quality (OQAQ score of 7) were both written by authors also involved in Cochrane reviews.

Limitations

We used the OQAQ to assess the quality of systematic reviews. One limitation of this checklist is that it does not distinguish between flaws in methodology (leading to bias) and flaws in reporting. Especially reviews published in peer reviewed journals have limited space and often methodological details are not reported. Therefore, we do not recommend that the reader dismiss reviews that received OQAQ scores less than 5; however, caution is advised when interpreting their conclusions. Although the quality of a review can be high according to a checklist, its validity can be compromised by overlooking important details in the included trials.[60] This information on individual trial quality, clinical heterogeneity, definition of included population, outcomes used and if children were reported separately in mixed reviews should be weighted with any quality score.

Recommendations

For Future Trials

At present, the most important challenge is to make sure that future trials are of the highest methodological quality from the design stage to the reporting stage. We need to prevent that systematic reviews and clinical guidelines conclude that the results of a given trial are probably biased or cannot be translated into practice for other reasons after it is completed and published. This effort will save money, time and, most importantly, our patient resources. International scientific meetings could be used to discuss study designs and protocols with all stakeholders in the field, and define patient populations, potential effect modifiers and clinically important outcomes. In order to be able to combine the results of future trials appropriately, an internationally accepted definition of acute asthma and its severity is needed. Along with these definitions for asthma, a standard set of well defined clinically important outcomes should be used in future trials on the management of acute asthma.

Errors, research misconduct and bias in intervention research could be reduced markedly if ongoing initiatives to register all trials at their inception, ensuring public access to trial protocols and all data generated, become successful.

For Future Pediatric Systematic Reviews

Journals could improve the quality of systematic reviews they publish by providing authors and peer reviewers with clear reporting criteria and by encouraging authors to have the protocol checked by peer reviewers before the actual systematic review is performed – as the major general medical journals already do for clinical trials.[11] To enhance the correct interpretation of findings, sources of funding of the included trials and of the review itself should be clearly stated. Until there is clear evidence that there are no differences between adults and children in their response to acute asthma interventions, separate analyses need to be conducted. Clinical heterogeneity should always be discussed. Because pediatric reviews often include small numbers of studies, the currently used statistical test to detect heterogeneity is inappropriate. Future reviews should use I^2 to measure statistical heterogeneity.[61] Reviewers need guidance on how to cope with the various forms of heterogeneity in reviews. As the validity and usefulness of systematic reviews cannot solely be based on scoring lists like the OQAQ, pediatricians with both clinical and methodological experience will be needed to author systematic reviews.

Conclusions

The methodological quality of both Cochrane and journal reviews on the management of acute asthma in children seems good, with Cochrane reviews being more rigorous. Yet, their usefulness for clinical practice is hampered by a lack of clear definitions of included populations and clinical important health outcomes, and separate reporting on children in mixed reviews. A major threat to these reviews' internal and external validity is the insufficient identification and handling of clinical, methodological, and statistical heterogeneity.

Acknowledgement

We thank Helena Buitelaar, medical student, for her help with the searches of the systematic reviews.

Reference List

1 Mannino DM, Homa DM, Pertowski CA, Ashizawa A, Nixon LL, Johnson CA. Surveillance for asthma - United States, 1960-1995. Morbidity and Mortality Weekly Report. CDC Surveillance Summaries 1998;47:1-27.

2 Camargo CA Jr (on behalf of the MARC investigators). Acute asthma among children presenting to the emergency department: The Multicentre Asthma Research Collaboration. Acad Emerg Med 1998;5:380 [abstract].

3 Camargo CA Jr (on behalf of the MARC investigators). Management of acute asthma in US emergency departments: The Multicenter Asthma Research Collaboration. American Journal of Respiratory and Critical Care Medicine 1998;157 [abstract].

4 Ian Chalmers, Matthias Egger, George Davey-Smith, Douglas G.Altman. Systematic Reviews in Health Care: Meta-analysis in Context. BMJ Publishing Group; 2001.

5 Greenhalgh T. Papers that summarise other papers (systematic reviews and meta-analyses). BMJ 1997 Sep 13;315(7109):672-5.

6 Assendelft WJ, Koes BW, Knipschild PG, Bouter LM. The relationship between methodological quality and conclusions in reviews of spinal manipulation. JAMA 1995 Dec 27;274(24):1942-8.

7 Choi PT, Halpern SH, Malik N, Jadad AR, Tramer MR, Walder B. Examining the evidence in anesthesia literature: a critical appraisal of systematic reviews. Anesth Analg 2001 Mar;92(3):700-9.

8 Delaney A, Bagshaw SM, Ferland A, Manns B, Laupland KB, Doig CJ. A systematic evaluation of the quality of meta-analyses in the critical care literature. Crit Care 2005 Oct 5;9(5):R575-R582.

9 Dixon E, Hameed M, Sutherland F, Cook DJ, Doig C. Evaluating meta-analyses in the general surgical literature: a critical appraisal. Ann Surg 2005 Mar;241(3):450-9.

10 Jadad AR, McQuay HJ. Meta-analyses to evaluate analgesic interventions: a systematic qualitative review of their methodology. J Clin Epidemiol 1996 Feb;49(2):235-43.

11 Jadad AR, Moher M, Browman GP, Booker L, Sigouin C, Fuentes M, et al. Systematic reviews and meta-analyses on treatment of asthma: critical evaluation. BMJ 2000 Feb 26;320(7234):537-40.

12 Kelly KD, Travers A, Dorgan M, Slater L, Rowe BH. Evaluating the quality of systematic reviews in the emergency medicine literature. Ann Emerg Med 2001 Nov;38(5):518-26.

13 Smyth RL, O'Byrne PM. The promise of and problems with evidence-based medicine for paediatric asthma management. Allergy 2002;57 Suppl 74:5-14.:5-14.

14 Oxman AD, Guyatt GH. Validation of an index of the quality of review articles. J Clin Epidemiol 1991;44(11):1271-8.

15 Oxman AD, Guyatt GH, Singer J, Goldsmith CH, Hutchison BG, Milner RA, et al. Agreement among reviewers of review articles. J Clin Epidemiol 1991;44(1):91-8.

16 Edmonds ML, Camargo CA, Jr., Pollack CV, Jr., Rowe BH. The effectiveness of inhaled corticosteroids in the emergency department treatment of acute asthma: a meta-analysis. Ann Emerg Med 2002 Aug;40(2):145-54.

17 Edmonds ML, Camargo CA, Jr., Brenner BE, Rowe BH. Replacement of oral corticosteroids with inhaled corticosteroids in the treatment of acute asthma following emergency department discharge: a meta-analysis. Chest 2002 Jun;121(6):1798-805.

18 Plotnick LH, Ducharme FM. Should inhaled anticholinergics be added to beta2 agonists for treating acute childhood and adolescent asthma? A systematic review. BMJ 1998 Oct 10;317(7164):971-7.

19 Plotnick LH, Ducharme FM. Acute asthma in children and adolescents: should inhaled anticholinergics be added to beta(2)-agonists? Am J Respir Med 2003;2(2):109-15.

20 Manser R, Reid D, Abramson M. Corticosteroids for acute severe asthma in hospitalised patients. Cochrane Database Syst Rev 2001;(1):CD001740.

21 Streetman DD, Bhatt-Mehta V, Johnson CE. Management of acute, severe asthma in children. Ann Pharmacother 2002 Jul;36(7-8):1249-60.

22 Aaron SD. The use of ipratropium bromide for the management of acute asthma exacerbation in adults and children: a systematic review. J Asthma 2001 Oct;38(7):521-30.

23 Amirav I, Newhouse MT. Metered-dose inhaler accessory devices in acute asthma: efficacy and comparison with nebulizers: a literature review. Arch Pediatr Adolesc Med 1997 Sep;151(9):876-82.

24 Blitz M, Blitz S, Beasely R, Diner BM, Hughes R, Knopp JA, et al. Inhaled magnesium sulfate in the treatment of acute asthma. Cochrane Database Syst Rev 2005; (4): CD003898.

25 Camargo CA, Jr., Spooner CH, Rowe BH. Continuous versus intermittent beta-agonists in the treatment of acute asthma. Cochrane Database Syst Rev 2003;(4):CD001115.

26 Castro-Rodriguez JA, Rodrigo GJ. beta-agonists through metered-dose inhaler with valved holding chamber versus nebulizer for acute exacerbation of wheezing or asthma in children under 5 years of age: a systematic review with meta-analysis. J Pediatr 2004 Aug;145(2):172-7.

27 Cates CJ, Crilly JA, Rowe BH. Holding chambers (spacers) versus nebulisers for beta-agonist treatment of acute asthma. Cochrane Database Syst Rev 2006; (2):CD000052.

28 Cheuk DK, Chau TC, Lee SL. A meta-analysis on intravenous magnesium sulphate for treating acute asthma. Arch Dis Child 2005 Jan;90(1):74-7.

29 Edmonds ML, Camargo CA, Saunders LD, Brenner BE, Rowe BH. Inhaled steroids in acute asthma following emergency department discharge. Cochrane Database Syst Rev 2000;(3):CD002316.

30 Edmonds ML, Camargo CA, Jr., Pollack CV, Jr., Rowe BH. Early use of inhaled corticosteroids in the emergency department treatment of acute asthma. Cochrane Database Syst Rev 2003;(3):CD002308.

31 Goodman DC, Littenberg B, O'Connor GT, Brooks JG. Theophylline in acute childhood

asthma: a meta-analysis of its efficacy. Pediatr Pulmonol 1996 Apr;21(4):211-8.

32 Graham V, Lasserson T, Rowe BH. Antibiotics for acute asthma. Cochrane Database Syst Rev 2001;(3):CD002741.

33 Ho AM, Lee A, Karmakar MK, Dion PW, Chung DC, Contardi LH. Heliox vs air-oxygen mixtures for the treatment of patients with acute asthma: a systematic overview. Chest 2003 Mar;123(3):882-90.

34 Mitra A, Bassler D, Goodman K, Lasserson TJ, Ducharme FM. Intravenous aminophylline for acute severe asthma in children over two years receiving inhaled bronchodilators. Cochrane Database Syst Rev 2005 Apr 18;(2):CD001276.

35 Osmond MH, Klassen TP. Efficacy of ipratropium bromide in acute childhood asthma: a meta-analysis. Acad Emerg Med 1995 Jul;2(7):651-6.

36 Plotnick LH, Ducharme FM. Combined inhaled anticholinergics and beta2-agonists for initial treatment of acute asthma in children. Cochrane Database Syst Rev 2000;(4): CD000060.

37 Rodrigo G, Pollack C, Rodrigo C, Rowe BH. Heliox for nonintubated acute asthma patients. Cochrane Database Syst Rev 2003;(4):CD002884.

38 Rodrigo GJ, Castro-Rodriguez JA. Anticholinergics in the treatment of children and adults with acute asthma: a systematic review with meta-analysis. Thorax 2005 Sep; 60(9):740-6.

39 Rowe BH, Keller JL, Oxman AD. Effectiveness of steroid therapy in acute exacerbations of asthma: a meta-analysis. Am J Emerg Med 1992 Jul;10(4):301-10.

40 Rowe BH, Bretzlaff JA, Bourdon C, Bota GW, Camargo CA, Jr. Magnesium sulfate for treating exacerbations of acute asthma in the emergency department. Cochrane Database Syst Rev 2000;(2):CD001490.

41 Rowe BH, Spooner C, Ducharme FM, Bretzlaff JA, Bota GW. Early emergency department treatment of acute asthma with systemic corticosteroids. Cochrane Database Syst Rev 2001;(1):CD002178.

42 Rowe BH, Spooner CH, Ducharme FM, Bretzlaff JA, Bota GW. Corticosteroids for preventing relapse following acute exacerbations of asthma. Cochrane Database Syst Rev 2001;(1):CD000195.

43 Smith M, Iqbal S, Elliott TM, Everard M, Rowe BH. Corticosteroids for hospitalised children with acute asthma. Cochrane Database Syst Rev 2003;(2):CD002886.

44 Travers A, Jones AP, Kelly K, Barker SJ, Camargo CA, Rowe BH. Intravenous beta2-agonists for acute asthma in the emergency department. Cochrane Database Syst Rev 2001;(2):CD002988.

45 The Cochrane Collaboration. Cochrane Handbook for Systematic Reviews of Interventions 4.2.5. 2005.

46 Chan AW, Hrobjartsson A, Haahr MT, Gotzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. JAMA 2004 May 26;291(20):2457-65.

47 Williamson PR, Gamble C, Altman DG, Hutton JL. Outcome selection bias in meta-analysis. Stat Methods Med Res 2005 Oct;14(5):515-24.

48 Sutton AJ, Duval SJ, Tweedie RL, Abrams KR, Jones DR. Empirical assessment of effect of publication bias on meta-analyses. BMJ 2000 Jun 10;320(7249):1574-7.

49 Higgins JP, Green S e. Cochrane Handbook for Systematic Reviews of Interventions 4.2.5 [updated May 2005]. http:/www.cochrane.org/resources/handbook/hbook.htm. 2005.

50 Moher D, Pham B, Jones A, Cook DJ, Jadad AR, Moher M, et al. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? Lancet 1998 Aug 22;352(9128):609-13.

51 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 Feb 1;273(5):408-12.

52 Higgins J, Thompson S, Deeks J, Altman D. Statistical heterogeneity in systematic reviews of clinical trials: a critical appraisal of guidelines and practice. J Health Serv Res Policy 2002 Jan;7(1):51-61.

53 Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002 Jun 15;21(11):1539-58.

54 Als-Nielsen B, Chen W, Gluud C, Kjaergard LL. Association of funding and conclusions in randomized drug trials: a reflection of treatment effect or adverse events? JAMA 2003 Aug;290(7):921-8.

55 Jorgensen AW, Hilden J, Gotzsche PC. Cochrane reviews compared with industry supported meta-analyses and other meta-analyses of the same drugs: systematic review. BMJ 2006 Oct 14;333(7572):782.

56 Barnes DE, Bero LA. Why review articles on the health effects of passive smoking reach different conclusions. JAMA 1998;279(19):1566-70.

57 Cramer K, Wiebe N, Moyer V, Hartling L, Williams K, Swingler G, et al. Children in reviews: methodological issues in child-relevant evidence syntheses. BMC Pediatr 2005 Sep 21;5:38.:38.

58 Bisgaard H. Long-acting beta(2)-agonists in management of childhood asthma: A critical review of the literature. Pediatr Pulmonol 2000 Mar;29(3):221-34.

59 Jadad AR, Cook DJ, Jones A, Klassen TP, Tugwell P, Moher M, et al. Methodology and reports of systematic reviews and meta-analyses: a comparison of Cochrane reviews with articles published in paper-based journals. JAMA 1998 Jul 15;280(3):278-80.

60 Hopayian K. The need for caution in interpreting high quality systematic reviews. BMJ 2001 Sep 22;323(7314):681-4.

61 Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003 Sep 6;327(7414):557-60.

CHAPTER 5

Fluid resuscitation in neonatal and pediatric hypovolemic shock: a Dutch Pediatric Society evidence-based clinical practice guideline

Nicole Boluyt, Casper W. Bollen, Albert P. Bos, Joke H. Kok, Martin Offringa

Intensive Care Med. 2006 Jul;32(7):995-1003

ABSTRACT

Objective To develop a clinical practice guideline that provides recommendations for the fluid, i.e. colloid or crystalloid, used for resuscitation in critically ill neonates and children up to the age of 18 years with hypovolemia.

Methods The guideline was developed through a comprehensive search and analysis of the pediatric literature. Recommendations were formulated by a national multidisciplinary committee involving all stakeholders in neonatal and pediatric intensive care and were based on research evidence from the literature and, in areas where the evidence was insufficient or lacking, on consensus after discussions in the committee.

Results Because of the lack of evidence in neonates and children, trials conducted in adults were considered. We found several recent meta-analyses that show excess mortality in albumin-treated groups, compared with crystalloid-treated groups, and one recent large randomized controlled trial that found evidence of no mortality difference. We found no evidence that synthetic colloids are superior to crystalloid solutions.

Conclusions Given the state of the evidence and taking all other considerations into account, the guideline-developing group and the multidisciplinary committee recommend that in neonates and children with hypovolemia the first-choice fluid for resuscitation should be isotonic saline.

Introduction

Hypovolemia is the most common cause of circulatory failure in children. When inadequate tissue perfusion is not recognized and treated during a narrow window of opportunity, critical tissue hypoxia may develop, leading to a cascade of events resulting in multiple organ failure and death. For this reason the concept of early goal-directed therapy was introduced by Rivers.[1] The authors show that this concept provides significant benefits with respect to outcome in adult patients.

The first step in the treatment of hypovolemic shock is adequate fluid resuscitation with either a crystalloid or a colloid solution. Pediatric advanced life-support guidelines recommend up to 60 ml/kg fluid resuscitation during treatment of hypovolemic and septic shock.[2] For decades there has been controversy over the relative benefits of crystalloid versus colloid solutions for fluid resuscitation of hypovolemic patients.[3] Since the publication of two systematic reviews in 1998 in the *British Medical Journal*[4,5] and the Cochrane Library,[6,7] the debate has intensified. These reviews of clinical trials conducted predominantly in adults consistently show an excess mortality of around 6% in critically ill patients who received human albumin solutions in comparison with patients who received a crystalloid. In the Netherlands, the current uncertainty has resulted in the plea for a national evidence-based guideline in the pediatric age group. The goal of this guideline is to define the current optimal choice of fluid used for treatment of neonates and children with circulatory failure due to hypovolemia, and to attain more uniformity in clinical practice. Recently the final version of this guideline was issued; it is now recommended and endorsed by the Dutch Pediatric Society.

In this paper we describe the methods and results of this guideline developmental process, including (1) a survey questionnaire to assess pre-guideline volume replacement strategies on neonatal and pediatric intensive care units in the Netherlands; (2) a systematic review of all randomized controlled trials on fluid resuscitation in hypovolemic neonates and children; (3) all systematic reviews on fluid resuscitation in hypovolemic adults; (4) other considerations taken into account to reach consensus in our national committee and (5) the final guideline recommendations.

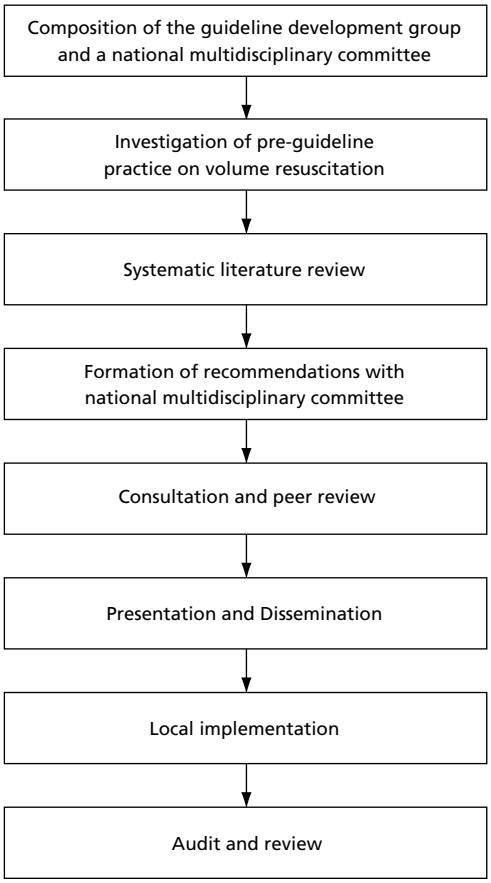
Methods

Figure 1 outlines the guideline development process.

To develop this clinical practice guideline, we formed a guideline-development group consisting of five members. A national multidisciplinary committee was formed comprising 29 members of all relevant disciplines and stakeholders (see Appendix).

As one of the goals of this guideline was to achieve more uniformity in clinical practice policy, we developed a questionnaire to first investigate the current Dutch

Figure 1 Guideline development process



pediatric practice of fluid solutions used for volume resuscitation. The questionnaire was sent to all directors of neonatal (n=10) and pediatric intensive care units (n=8) in the Netherlands.

Based on the results of this questionnaire, the following questions were formulated: (1) What type of fluid solution should be used for initial resuscitation of hypovolemia in neonates and children? (2) What is the optimal amount of fluid to be given and at what infusion rate? (3) What are the possible side effects related to each type of fluid, such as hypernatremia and peripheral edema?

Studies were identified by sensitive computerized searches of Medline (1966-2000), Embase (1988-2000) and the Cochrane Library, with the help of a clinical librarian. In addition, reference lists of all available articles were reviewed to identify additional citations not found in the computerized search. Studies written in English, French, German and Dutch were eligible for inclusion. We searched for guidelines, systematic

reviews, meta-analyses and randomized controlled trials on volume resuscitation in critically ill neonates, children and adults with hypovolemia due to septic shock, trauma, dehydration, hemorrhage and post-cardiac surgery. With regard to question 2 we searched for studies comparing different volumes and rates of infusion in critically ill children. In October 2005 we repeated our literature search and looked for additional systematic reviews and randomized controlled trials in neonates and children. Studies were included only when they concerned clinically relevant outcomes i.e. mortality and major morbidity such as pulmonary edema, length of stay in hospital, intraventricular hemorrhage or impaired neurodevelopment. Further details on the search strategy can be requested from the authors.

Each study was assessed independently for its methodological quality by two investigators using critical appraisal forms originally published in *JAMA* (http://ugi.usersguides.org/usersguides/hg/hh_start.asp). Disagreement between these raters was resolved by consensus. Each article was assigned a 'level of evidence' (Table 1), which in turn influenced the 'grade of recommendation' (Table 2). These grades of recommendation originate from the US Agency for Health Care Policy and Research.[8]

Table 1 Classification of evidence levels

1a	Evidence obtained from meta-analysis of randomized controlled trials
1b	Evidence obtained from at least one randomized controlled trial
2a	Evidence obtained from at least one well-designed controlled study without randomization
2b	Evidence obtained from at least one other type of well-designed quasi-experimental study*
3	Evidence obtained from well-designed non-experimental descriptive studies, such as comparative studies, correlation studies and case studies
4	Evidence obtained from expert committee reports or opinions and/or clinical experiences of respected authorities

* Refers to a situation in which implementation of an intervention is outside of the control of the investigators, but an opportunity exists to evaluate its effect

Table 2 Classification of grades of recommendations

A	Requires at least one randomized controlled trial as part of a body of literature of overall good quality and consistency addressing specific recommendation (Evidence levels 1a, 1b)
B	Requires the availability of well conducted clinical studies but no randomized clinica trials on the topic of recommendation (Evidence levels 2a, 2b, 3)
C	Requires evidence obtained from expert committee reports or opinions and/or clinical experiences of respected authorities. Indicates an absence of directly applicable clinical studies of good quality (Evidence level 4)

Note: These classifications of types of evidence and the corresponding grades of recommendation originate from the US Agency for Health Care Policy and Research[8]

The committee met on two separate occasions to formulate the final recommendations. In the event that there was not enough evidence or the evidence was of poor quality, the recommendations were based on consensus after discussion in the committee and finally by show of hands. The following 'other considerations' for reaching consensus were taken into account: (a) potential side effects of colloids and crystalloids, (b) current insight in pathophysiological mechanisms and their impact on the applicability of evidence from adults to children and neonates and (c) costs. All three aspects will be discussed below.

After a draft guideline was released it was piloted among end-users in the national multidisciplinary committee's hospital wards; feedback was received and included in the final version of the guideline. A comprehensive technical report can be obtained from the authors.

Results

Questionnaire

The response to the questionnaire was 10/10 (100%) of the neonatologists and 7/8 (88%) of the pediatric intensivists. First-choice fluid for volume resuscitation was in 50% a crystalloid and in 50% a colloid solution for both neonatologists and pediatric intensivists. The neonatologists used human albumin as a priority colloid, and the pediatric intensivists predominantly used a synthetic colloid, e.g. Gelofusine.

Literature

We did not identify any evidence-based guidelines on this topic. Our search for systematic reviews, meta-analyses and randomized controlled trials identified 93 citations, of which 65 met the inclusion criteria. Twelve articles concerned children or neonates. Included systematic reviews and randomized trials are summarized in Table 3 and Table 4, respectively.

Question 1: What type of fluid solution should be used for initial resuscitation of hypovolemia?

Premature and full term neonates

We identified one meta-analysis by Kirpalani, which was excluded because it does not analyze hypovolemic neonates separately.[9] We identified five randomized controlled trials.[10-14] Four studies did not meet our inclusion criteria: three focused on giving fluids prophylactically after birth[10-12] and one on giving albumin when there was hypoalbuminemia.[13] Only the study by So[14] met our inclusion criteria. The investigators undertook a randomized controlled trial to study the efficacy of a colloid (i.e. 5% albumin)

Table 3 Summary of included systematic reviews comparing the effect of any fluid solution with other fluid solutions

Study	Year	Age group	Inter-vention	Comparison	Outcomes	Adequate search strategy	Quality assessment of RCT's	Description of articles	Focused clinical question	Description of selection procedure of RCT's	Description of data extraction	Valid statistical pooling	Assessment of hetero-geneity	RR (95% CI)[a]	Level of evidence
Velanovich[24]	1989	Adult	Colloid	Crystalloid	Death	No	No	Yes	Yes	No	No	Yes	No	NG	NA
Bisonni[21]	1991	Adult	Colloid	Crystalloid	Death	No	No	No	Yes	No	No	No	No	NG	NA
Schierhout[5]	1998	Adult	Colloid	Crystalloid	Death	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	1.52 (1.08-2.13)	1a
Choi[23]	1999	Adult	Colloid	Crystalloid	Death, pulmonary edema	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	0.86 (0.63-1.17)	1a
Bunn[22]	2000	Adult	Colloid	Colloid	Death	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	0.99 (0.69-1.42)	1a
Alderson[7]	2002	Adult	Albumin	No albumin	Death	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	1.46 (0.97-2.22)	1a

NG, not given; NA, not applicable

a Mortality RR in hypovolemic patients

Table 4 Summary of included pediatric randomized trials comparing the effect of any fluid solution with other fluid solutions, and the recent SAFE trial

Study	Year	Age group	No of patients	Inter-vention	Compari-son	Outcome	Adequate Randomi-zation	Sufficient follow up	Blinding of patients	Blinding of treatment	Blinding of effect assessment	Patients prognostic equal	Equal concomitant treatment	RR (95% CI) mortality	Level of evidence
Boldt[15]	1993	Child	30	Albumin	HES	Death, pulmonary edema	Yes	No	No	No	No	No	No	No deaths and no pulmonary edema occurred	1b
So[14]	1997	Child	63	Albumin	Crystalloid	Death, CLD,* IVH**	Yes	Yes	No	No	Yes	Yes	Yes	1.36 (0.69-2.66)	1b
SAFE Study Investigators[26]	2004	Adult	6997	Albumin	Isotonic saline	Death	Yes	Yes	Yes	Yes	Yes	Yes	Yes	0.99 (0.91-1.09)	1b
Wills[20]	2005	Child	512	Ringer's lactate or 6% dextran 70 (a colloid) or 6% hydroxyethyl starch (a colloid)		No. of days in hospital	Yes	Yes	Yes	Yes	Yes	Yes	Not stated	Not given	1b

* CLD= chronic lung disease

** IVH= intraventricular hemorrhage

versus a crystalloid (i.e. isotonic saline) in the treatment of 63 hypotensive preterm neonates. Outcomes, as assessed by the number of infants that died (RR 1.36; 95% CI 0.69-2.66), chronic lung disease (RR 0.48; 95% CI 0.13-1.87) or intraventricular hemorrhage (RR 1.52; 95% CI 0.91-4.87), did not differ significantly between the groups, but the wide confidence intervals indicate that the study was underpowered.

Older children

Our search identified six randomized trials.[15-20] Four studies were excluded: one of them assessed albumin supplementation for hypoalbuminemia,[16] another assessed hypertonic saline in neurotrauma,[18] the third study included children with severe malaria and metabolic acidosis, not hypovolemia per se,[19] and the study by Ngo[17] used surrogate endpoints. Together with the recently published Wills trial,[20] the Ngo study is the only large randomized controlled trial in children. We will briefly discuss the results of this trial, although it does not meet our inclusion criterion on clinical relevant outcomes.

Ngo performed a randomized, double blind trial comparing the efficacy of four different fluid regimens (dextran 70, 3% gelatin, lactated Ringer's, and isotonic saline) in the initial management of dengue shock syndrome (DSS) in 222 children aged 1-15 years. The primary outcome measures were the initial pulse pressure recovery time and the occurrence of subsequent episodes of shock. Secondary outcome measures included the development of 'any complication' of fluid therapy. There were no deaths in any of the groups and there were no differences in the 'reshock' rate among the four groups. Six children had allergic reactions after colloid therapy (five received gelatin and one dextran), defined as fever and chills. One child in the gelatin group had severe epistaxis and another child in the dextran group a large hematoma at a site of minor trauma. In this study no clear benefits of any one of the four fluids in improving these surrogate endpoints could be demonstrated. Normal saline performed as well as the colloid solutions.

Boldt conducted a randomized study in 30 children less than 3 years of age undergoing cardiac surgery.[15] They were randomly assigned to receive albumin or low-molecular-weight hydroxyethyl starch solution (6% LMW-HES). No deaths occurred and none of the children had signs of pulmonary edema, but the wide confidence intervals indicate that the study was underpowered.

Recently, Wills conducted a large randomized controlled trial in 512 children with dengue shock syndrome, comparing three resuscitation fluids; Ringer's lactate, 6% dextran 70 and 6% hydroxyethyl starch.[20] A total of 383 children with moderately severe shock were randomly assigned to receive one of the three fluids and 129 children with severe shock to receive one of the colloids. The primary outcome was the requirement for supplemental intervention with rescue colloid. Secondary outcome measures were 'time taken to achieve initial and sustained cardiovascular stability' and 'number of days in hospital'. The relative risk of a requirement for rescue colloid was 1.08 (95% CI 0.78-1.47) among children with moderate shock who received Ringer's lac-

tate as opposed to either of the colloid solutions. There were no differences in the time to final cardiovascular stability or the number of days in the hospital.

Adults

Because of the lack of randomized controlled trials in children, we included trials conducted in critically ill adults. We found seven meta-analyses.[4-7,21-25] The systematic review by Wilkes was excluded because it did not analyze hypovolemic patients separately.[25] Recently a large randomized controlled trial, the Saline versus Albumin Fluid Evaluation (SAFE) Study, was published.[26] Because of poor methodological quality we will not discuss the two older systematic reviews that compare colloids with crystalloids in fluid resuscitation here.[21,24] Instead, the results of the remaining four reviews and the SAFE Study will be discussed below.

All four remaining systematic reviews are of good methodological quality (Table 3). The reviews by Alderson and Schierhout comparing colloids with crystalloids both show a 6% increase in mortality in the group receiving albumin.[5,6]

The meta-analysis by Choi[23] observed no difference between crystalloid and colloid resuscitation with respect to pulmonary edema [pooled RR 0.84 (0.25-2.45)] and mortality [pooled RR 0.86 (0.63-1.17)]. However, the power of the aggregated data was insufficient to detect small but potentially clinically important differences. Subgroup analysis showed a statistically significant reduction in mortality rate in trauma in favor of crystalloid resuscitation (RR 0.39, 95% CI 0.17-0.89).

The meta-analysis by Bunn did not show evidence that one colloid solution is more effective or safe than any other: albumin vs. HES: RR 1.17 (0.91-1.50), albumin vs. gelatin: RR 0.99 (0.69- 1.42).[22] Again, in this study the confidence intervals are wide and the results do not exclude clinically relevant differences among colloids.

The SAFE Study investigators conducted a large randomized controlled trial in adult patients who had been admitted to the ICU and required fluid administration to maintain or increase intravascular volume.[26] A total of 6,997 patients were randomly assigned to receive 4% albumin or isotonic saline. The relative mortality risk among patients assigned to receive albumin compared with those assigned saline was 0.99 (95% CI 0.91-1.09). The authors conclude that there is evidence of no difference in mortality rate.

It is possible that there are certain subgroups where either colloids or crystalloids are more effective. These four meta-analyses comprised different patient categories, namely trauma, hypoalbuminemia, hypovolemia, sepsis, burns, and cardiopulmonary surgery. There was no evidence of advantage of colloids over crystalloids in any of these different indication subgroups, although in most cases the results were inconclusive because of lack of statistical power. Especially in patients after trauma or burns, albumin appeared to be associated with increased mortality and increased ventilator requirement. In trauma patients, the relative mortality rates when comparing albumin

and saline were between 0.98 and 5.88. The SAFE Study included post-hoc subgroup analyses. The relative risk of death in the albumin group compared with patients in the saline group was 1.36 (95% CI 0.99-1.87) for trauma patients, 0.87 (95% CI 0.74-1.02) for patients with sepsis and 0.93 (95% CI 0.61-1.41) for patients with respiratory distress syndrome. It must, however, be noted that in large studies such subgroup differences frequently occur by chance, and common wisdom holds that hypotheses generated in this manner should be evaluated in specifically designed and appropriately powered future studies.

Question 2: What is the optimal amount to give and at what infusion rate?

We identified only one study by Carcillo about the role of early and rapid fluid resuscitation in children with septic shock.[27] Included were 34 children (median age 13.5 months) with septic shock who all required vasopressor and/or inotropic support. Therapeutic decisions were left to the attending staff in this observational study. At 1 h and 6 h, respectively, group1 (n=14) received 11±6 and 71±29 ml/kg (mean ± SD) of fluid; group 2 (n=11) received 32±5 and 108±54 ml/kg of fluid; and group 3 (n=9) received 69±19 and 117±29 ml/kg of fluid. Fluids used were 5% albumin, fresh frozen plasma, cryoprecipitate, isotonic saline and lactated Ringer's solution. Details on which patient received which fluid were not given in the paper. Survival in group 3 (8 of 9 patients) was significantly better than in group 1 (6 of 14 patients) or group 2 (4 of 11 patients). The authors concluded that rapid fluid resuscitation in excess of 40 ml/kg in the first hour following emergency department presentation of children with septic shock is associated with improved survival. However, since the treatment groups were assigned non-randomly and the choice of treatment was based on clinical criteria that were determined by individual physicians, and since no adjustments were made for co-interventions like the use of inotropics, these observations cannot prove cause and effect. *Level of evidence: 2b.*

There is no evidence in children and neonates with hypovolemia not caused by septic shock about the optimal volume to be used and the velocity of fluid resuscitation. The efficacy of fluid replacement depends on the existing microvascular pressures, the compliance of the interstitial space and the permeability of the microvascular barrier. Hence both the volume and velocity of fluid resuscitation should be determined individually.

Question 3: What are the possible side effects related to the type of fluid used, such as hypernatremia and peripheral edema?

Hypernatremia

Hypernatremia is thought by some to play a role in the development of intraventricular hemorrhage in neonates. There is a concern that by giving isotonic saline hyperna-

tremia will be induced, because of a diminished renal sodium clearance potential in the first week of life. One of the outcomes assessed in the study by So was the serum sodium concentration.[14] The investigators included 63 hypotensive premature infants, randomly assigned to receiving a colloid or crystalloid solution. They could not find a significant difference in mean sodium concentration, or the rate of intraventricular hemorrhage or mortality, between the two groups. This lack of difference observed may be partly due to small numbers of patients. Furthermore, in most countries 5% albumin contains almost the same amount of sodium as isotonic saline (145 mmol/l vs.154 mmol/l).

Peripheral edema

Both the use of colloids and crystalloids is accompanied by the occurrence of peripheral edema.[16,28-31] More important to know is how quickly such edema can be mobilized and whether there are any clinically relevant consequences. In our sensitive searches for evidence, no studies that give answers to these questions could be found.

Other considerations taken into account at the consensus meeting

Pathophysiology and applicability of the evidence from adults in neonates and children

If the alveolo-capillary membrane is intact, the lungs are well protected against a drop of colloid osmotic pressure (or hypoalbuminemia); if the membrane is damaged, the infusion of colloid aimed at increasing the colloid osmotic pressure is illusive, since the colloids leak into the interstitium and could even amplify the pulmonary and peripheral edema.

Although most of the evidence in this field comes from studies in critically ill adults, the results of these investigations do not differ from those studies that have been conducted in critically ill neonates and children; there is just more information on adults. All studies fail to show an advantage of colloids over crystalloids when survival rates are considered. In addition, there is no clear evidence whether colloids or crystalloids confer benefit in certain subgroups of patients with shock, i.e. septic, hemorrhagic or hypovolemic. The reported increased mortality rate due to the use of albumin could be explained by the distribution of albumin across the capillary membrane, a process that is accelerated in critically ill patients.[54] Increased leakage of colloids into the extra vascular spaces might reduce the oncotic pressure difference across the capillary wall, making edema more likely. Yet, given the results of the SAFE Study, this mechanism may exist, but it may be less important in clinical practice than we thought previously. Although the distribution of body fluids in the neonate differs from that in adults, there is currently no reason to believe – either from pathophysiological theory or empirical evidence – that once a capillary leak exists, albumin would be beneficial in children.

It is, however, possible that in contrast to albumin, which is a small molecule (60 kDa), synthetic colloids with larger molecules (HES, 200 kDa) do not leak into the interstitial space. Yet, so far there is no evidence that some colloids are more effective or safer than others, when clinically relevant outcomes are considered.

Isotonic saline is distributed equally throughout the extra cellular space. Because the extra cellular fluid is one-fourth intravascular and three-fourths interstitial, only one-fourth of the infusate remains intravascular. For this reason some members of the multidisciplinary consensus group believed that it might be more efficient to use a synthetic colloid after initial crystalloids in patients with refractory hypovolemia and significant hemodynamic problems (e.g. those with severe sepsis). However, the SAFE Study showed that the ratio of the volume needed to maintain stable circulation of albumin to the same volume of isotonic saline administered was only 1.4. Again, we currently have no reason to believe that this ratio would be different in neonates and children.

Costs

Colloid solutions are much more expensive than crystalloid solutions: 1 l of albumin currently costs around 140 Euro (152 US$), 1 l of HES costs 25 Euro (27 US$) and 1 l of isotonic saline costs 1.5 Euro (1.6 US$).

The committee's recommendations

As colloids are biological products with a potential infection hazard or a risk of anaphylactic reaction and because they are much more expensive than crystalloids, it was felt by the national multidisciplinary committee that their benefits over crystalloids should be proved before they were used. Given the state of the evidence and taking all other considerations into account, the guideline-developing group and the national multidisciplinary committee recommend that in neonates and children with hypovolemia the first-choice fluid for initial resuscitation should be isotonic saline. When large amounts of fluids are required (e.g. in sepsis), it is possible to use a synthetic colloid because of its longer duration in the circulation. The initial fluid volume should be 10-20 ml/kg, with repeated doses based on individual clinical response (Table 5).

Table 5 The guideline's recommendations

1	In neonates and children with hypovolemia the first choice fluid for initial resuscitation is isotonic saline (Grade A)*
2	When large amounts of fluids are required (e.g. sepsis), it is possible to use a synthetic colloid because of its longer duration in the circulation (Grade C)
3	The initial fluid volume should be 10-20 ml/kg and repeated doses should be based on individual clinical response (Grade C)

* In adults there is grade A evidence. For reasons described in this paper this evidence is thought to be applicable to neonates and children.

Discussion

In adult patients the current volume of evidence on the solutions of choice for fluid resuscitation of hypovolemic patients is large, consistently showing no advantage of colloids over crystalloids. Despite this, several surveys have shown that this evidence has not changed clinical practice: the majority of physicians still use colloid products.[33-35] More interestingly, most physicians could not state reasons for choosing between products.[34] To bridge this apparent gap between the evidence and actual clinical practice, guidelines may be needed.[36,37] Clinical practice guidelines are seen as powerful tools to achieve effective and efficient care, but have been demonstrated to be effective only if there is sufficient rigor in their method of development.[38] Having completed a rigorous and objective synthesis of the evidence base, the guideline-development group must make what is essentially a subjective judgment on the recommendations that can validly be made on the basis of the available evidence.[39] Subjective judgment risks the reintroduction of bias into the process. However, in high-quality guideline-development processes this is not the judgment of one individual but of a carefully composed multidisciplinary group. An additional safeguard here is the requirement for the guideline-development group to present clearly the evidence on which the recommendation is based, and to make the link between the evidence and the recommendations explicit, explaining how the evidence was interpreted.[39] Thus, the summarized research evidence is only one factor in clinical guidelines, and the 'other considerations' are just as important in reaching consensus on the recommendations. This holds especially when there is insufficient research evidence. Therefore, these 'other considerations' should be clearly presented in any guideline or its technical report.[40]

Before the present guideline was developed there was controversy in the Netherlands about the fluid of choice for volume replacement; about 50% of neonatologists and pediatric intensivists used colloids. We found that in children the volume and quality of the available research evidence is limited. The final recommendations are largely based on the 'other considerations', in this case the potential side effects of colloids and crystalloids, current insight into pathophysiological mechanisms and their impact on the applicability of evidence from adults to children and neonates, and costs. We involved 29 stakeholders in Dutch neonatal and pediatric intensive care practice to formulate the recommendations and unanimously decided that isotonic saline should be the first choice, because it is equally effective, safe and up to 100 times cheaper than albumin. We realize that in other settings and countries the 'other considerations' may play a different role and have a different bearing on decision making. As the process of development of this guideline was explicit, pediatric intensivists and neonatologists outside the Netherlands may make their own recommendations based on the information we have assembled.

We believe that the introduction of this nation-wide guideline for fluid resuscita-

tion may contribute to an optimal, cost-effective universal treatment strategy in pediatric patients with imminent circulatory failure.

Conclusion

Given the state of the evidence and taking other relevant considerations into account, the guideline-developing group and the Dutch national multidisciplinary committee recommend that in neonates and children with hypovolemia the first-choice fluid for resuscitation should be isotonic saline.

This guideline will be updated in July 2008

Acknowledgement

The work was funded by the 'Practice Guidelines Program' in the Academic Medical Centre in Amsterdam.

Appendix National Multidisciplinary Committee

Neonatologists

P. Andriessen, MD, W. Baerts, MD, PhD, P.L.J. Degraeuwe, MD, W.P.F. Fetter, MD, PhD, W.B. Geven, MD, PhD, A.H.L.C. van Kaam, MD, K.D. Liem, MD, PhD, D.W.E. Roofthooft, MD, M.G.A. Verboon-Maciolek, MD, F.J. Walther, MD PhD, F. Brus, MD, PhD, W.W.M. Hack, MD, PhD, J.W.F.M. Jacobs, MD.

Pediatric intensivists

P. Dahlem, MD, R.J.B.J. Gemke, MD, PhD, L.G.F.M. van 't Hek, MD, C. Buysse, MD, G.D. Vos, MD, A.J. van Vught, MD, PhD, W. de Weerd, MD, N. van der Lelij, MD, F.B. Plötz, MD, PhD, L. Veenhuizen, MD, W.B. Vreede, MD, C.M. Walhof, MD

Other

M. Hemmink, nurse neonatology, D. Tol, nurse intensive care
T.W.J. Schulpen, MD, PhD, Director Quality Office, Dutch Pediatric Society
C.R. Lincke, MD, PhD, Chair of the Society's Guideline Development Committee, Dutch Pediatric Society

None of the members of the study group or the national multidisciplinary committee had any connections with the pharmaceutical industry or had otherwise any financial interest in the final recommendations.

Reference List

1 Rivers E, Nguyen B, Havstad S, Ressler J, Muzzin A, Knoblich B, Peterson E, Tomlanovich M (2001) Early goal-directed therapy in the treatment of severe sepsis and septic shock. N Engl J Med 345:1368-1377

2 Ceneviva G, Paschall JA, Maffei F, Carcillo JA (1998) Hemodynamic support in fluid-refractory pediatric septic shock. Pediatrics 102:e19

3 Offringa M (1998) Excess mortality after human albumin administration in critically ill patients. Clinical and pathophysiological evidence suggests albumin is harmful. BMJ 317:223-224

4 Cochrane Injuries Group Albumin Reviewers (1998) Human albumin administration in critically ill patients: systematic review of randomised controlled trials. BMJ 317:235-240

5 Schierhout G, Roberts I (1998) Fluid resuscitation with colloid or crystalloid solutions in critically ill patients: a systematic review of randomised trials. BMJ 316:961-964

6 Alderson P, Schierhout G, Roberts I, Bunn F (2000) Colloids versus crystalloids for fluid resuscitation in critically ill patients. Cochrane Database Syst Rev:CD000567

7 Alderson P, Bunn F, Lefebvre C, Li WP, Li L, Roberts I, Schierhout G (2002) Human albumin solution for resuscitation and volume expansion in critically ill patients. Cochrane Database Syst Rev:CD001208

8 US Department of Health and Human Services. Agency for Health Care Policy and Research. Acute pain management: operative or medical procedures and trauma. Rockville (MD): The Agency; 1993. Clinical Practice Guideline No.1. AHCPR Publication No. 92-0023. p 107

9 Kirpalani H (2001) The use of albumin and/or colloids in newborns: A meta-analysis. Pediatr Crit Care Med 2:S14-S19

10 Northern Neonatal Nursing Initiative Trial Group (1996) Randomised trial of prophylactic early fresh-frozen plasma or gelatin or glucose in preterm babies: outcome at 2 years. Lancet 348:229-232

11 Northern Neonatal Nursing Initiative [NNNI] Trial Group (1996) A randomized trial comparing the effect of prophylactic intravenous fresh frozen plasma, gelatin or glucose on early mortality and morbidity in preterm babies. Eur J Pediatr 155:580-588

12 Bland RD, Clarke TL, Harden LB, Meyer JL, Ries JP, Madden WA, Crast FW, Coyer WF, Bass JW (1973) Early albumin infusion to infants at risk for respiratory distress. Arch Dis Child 48:800-805

13 Greenough A, Emery E, Hird MF, Gamsu HR (1993) Randomised controlled trial of albumin infusion in ill preterm infants. Eur J Pediatr 152:157-159

14 So KW, Fok TF, Ng PC, Wong WW, Cheung KL (1997) Randomised controlled trial of colloid or crystalloid in hypotensive preterm infants. Arch Dis Child Fetal Neonatal Ed 76:F43-F46

15 Boldt J, Knothe C, Schindler E, Hammermann H, Dapper F, Hempelmann G (1993) Volume replacement with hydroxyethyl starch solution in children. Br J Anaesth 70:661-665

16 Greenhalgh DG, Housinger TA, Kagan RJ, Rieman M, James L, Novak S, Farmer L, Warden GD (1995) Maintenance of serum albumin levels in pediatric burn patients: a prospective, randomized trial. J Trauma 39:67-73

17 Ngo NT, Cao XT, Kneen R, Wills B, Nguyen VM, Nguyen TQ, Chu VT, Nguyen TT, Simpson JA, Solomon T, White NJ, Farrar J (2001) Acute management of dengue shock syndrome: a randomized double-blind comparison of 4 intravenous fluid regimens in the first hour. Clin Infect Dis 32:204-213

18 Simma B, Burger R, Falk M, Sacher P, Fanconi S (1998) A prospective, randomized, and controlled study of fluid management in children with severe head injury: lactated Ringer's solution versus hypertonic saline. Crit Care Med 26:1265-1270

19 Maitland K, Pamba A, English M, Peshu N, Marsh K, Newton C, Levin M (2005) Randomized trial of volume expansion with albumin or saline in children with severe malaria: preliminary evidence of albumin benefit. Clin Infect Dis 40:538-545

20 Wills BA, Nguyen MD, Ha TL, Dong TH, Tran TN, Le TT, Tran VD, Nguyen TH, Nguyen VC, Stepniewska K, White NJ, Farrar JJ (2005) Comparison of three fluid solutions for resuscitation in dengue shock syndrome. N Engl J Med 353:877-889

21 Bisonni RS, Holtgrave DR, Lawler F, Marley DS (1991) Colloids versus crystalloids in fluid resuscitation: an analysis of randomized controlled trials. J Fam Pract 32:387-390

22 Bunn F, Alderson P, Hawkins V (2001) Colloid solutions for fluid resuscitation. Cochrane Database Syst Rev:CD001319

23 Choi PT, Yip G, Quinonez LG, Cook DJ (1999) Crystalloids vs. colloids in fluid resuscitation: a systematic review. Crit Care Med 27:200-210

24 Velanovich V (1989) Crystalloid versus colloid fluid resuscitation: a meta-analysis of mortality. Surgery 105:65-71

25 Wilkes MM, Navickis RJ (2001) Patient survival after human albumin administration. A meta-analysis of randomized, controlled trials. Ann Intern Med 135:149-164

26 Finfer S, Bellomo R, Boyce N, French J, Myburgh J, Norton R (2004) A comparison of albumin and saline for fluid resuscitation in the intensive care unit. N Engl J Med 350:2247-2256

27 Carcillo JA, Davis AL, Zaritsky A (1991) Role of early fluid resuscitation in pediatric septic shock. JAMA 266:1242-1245

28 Ernest D, Belzberg AS, Dodek PM (1999) Distribution of normal saline and 5% albumin infusions in septic patients. Crit Care Med 27:46-50

29 Lucas CE, Weaver D, Higgins RF, Ledgerwood AM, Johnson SD, Bouwman DL (1978) Effects of albumin versus non-albumin resuscitation on plasma volume and renal excretory function. J Trauma 18:564-570

30 Pockaj BA, Yang JC, Lotze MT, Lange JR, Spencer WF, Steinberg SM, Topalian SL, Schwartzentruber DJ, White DE, Rosenberg SA (1994) A prospective randomized trial

evaluating colloid versus crystalloid resuscitation in the treatment of the vascular leak syndrome associated with interleukin-2 therapy. J Immunother 15:22-28

31 Prien T, Backhaus N, Pelster F, Pircher W, Bunte H, Lawin P (1990) Effect of intraoperative fluid administration and colloid osmotic pressure on the formation of intestinal edema during gastrointestinal surgery. J Clin Anesth 2:317-323

32 Berger A (1998) Why albumin may not work. BMJ 317:240

33 Boldt J, Lenz M, Kumle B, Papsdorf M (1998) Volume replacement strategies on intensive care units: results from a postal survey. Intensive Care Med 24:147-151

34 Miletin MS, Stewart TE, Norton PG (2002) Influences on physicians' choices of intravenous colloids. Intensive Care Med 28:917-924

35 Soderlind M, Salvignol G, Izard P, Lonnqvist PA (2001) Use of albumin, blood transfusion and intraoperative glucose by APA and ADARPEF members: a postal survey. Paediatr Anaesth 11:685-689

36 Haines A, Jones R (1994) Implementing findings of research. BMJ 308:1488-1492

37 Woolf SH, Grol R, Hutchinson A, Eccles M, Grimshaw J (1999) Clinical guidelines: potential benefits, limitations, and harms of clinical guidelines. BMJ 20;318:527-530

38 Farquhar CM, Kofa EW, Slutsky JR (2002) Clinicians' attitudes to clinical practice guidelines: a systematic review. Med J Aust 177:502-506

39 Harbour R, Miller J (2001) A new system for grading recommendations in evidence based guidelines. BMJ 323:334-336

40 Burgers JS, van Everdingen JJ (2004) Beyond the evidence in clinical guidelines. Lancet 364:392-393

CHAPTER 6

Implementation and evaluation of a pediatric evidence-based guideline for fluid resuscitation in hypovolemia

Nicole Boluyt*, Merit M. Tabbers*, Martin Offringa

Submitted for publication

* Both authors contributed equally to the work presented

ABSTRACT

Objective To evaluate a systematic approach to the development and nationwide implementation of an evidence-based pediatric guideline on first choice fluid for resuscitation in hypovolemia.

Methods We investigated fluid prescribing behavior for hypovolemia at three stages: 1) before guideline development in 2000 2) after guideline development in 2004 and 3) after active implementation in a large sample of pediatric specialists in 2006, and identified potential barriers and facilitators for guideline implementation. In order to minimize costs and to optimize implementation effect, we continuously adjusted implementation strategies according to identified barriers. Implementation success was evaluated using both questionnaires and data from medical records.

Results A total of 144 specialists caring for children were surveyed. The most remarkable change in fluid prescribing behavior occurred after guideline development and dissemination: normal saline use by neonatologists increased from 22-89% (depending on hypovolemia type) to 100% for all hypovolemia indications, and by pediatric intensivists from 43-71% to 88-100%; synthetic colloid use by pediatric intensivists declined from 29-43% to 0-13% accompanied by a reduction in albumin use by neonatologists from 11-44% to zero. Data from medical records were in keeping with reported behavior.

Conclusions After active guideline implementation targeting identified barriers most of the Dutch pediatric specialist's management behavior was in concordance with the national guideline, i.e. they used normal saline. To successfully implement clinical guidelines and reduce the huge cost of active implementation, any guideline development should consider implementation right from the beginning and involve all stakeholders in the developmental process. Implementation strategies should be targeting identified barriers and will therefore always be guideline specific.

Introduction

Hypovolemia is the most common cause of circulatory failure in children. It may result in the development of multiple organ failure syndrome when not treated adequately with volume replacement.

For over 100 years now, controversy exists over whether colloid or crystalloid solutions should be used for volume therapy.[1] Several systematic reviews have shown that the use of albumin for fluid resuscitation in hypovolemic patients was associated with a less favorable outcome compared with non-albumin infusions and colloids are not superior to crystalloids.[2-6] Colloids are costly and have more side effects. The publication of these systematic reviews has sparked considerable debate among pediatricians and neonatologists,[7] as the reports included four neonatal studies showing a similar trend for less favorable outcome associated with the use of albumin infusions as in adults. The uncertainty surrounding the optimal fluid for management of hypovolemia has resulted in a Dutch national evidence-based guideline. The guideline recommends normal saline as the first choice fluid in hypovolemic neonates and children up to 18 years.[8]

Clinical practice guidelines can facilitate translation of research into clinical practice and are seen as powerful tools to achieve effective care, reduce variability in daily practice, and may reduce costs.[9] Unfortunately, it is well known that many guidelines are not used in daily practice unless they are actively implemented pursuing consolidation of behavior change.[10;11] Quality-of-care indicators are used to determine failure or success of guideline implementation.[12;13]

So far, there is little experience with the implementation of pediatric evidence-based guidelines. Our aim was to evaluate, both qualitatively and quantitatively, a systematic approach to the development and nationwide implementation of an evidence-based pediatric guideline on the first choice fluid for volume resuscitation in the hope that some important lessons can be learnt for future pediatric guidelines. In this paper we describe 1) pre- and post implementation questionnaire surveys to investigate first choice fluid use in various pediatric departments 2) implementation barriers and facilitators 3) the tailored implementation strategies that were used 4) quality-of-care indicators that we used to evaluate the implementation success, and 5) actual nationwide implementation success as measured by these indicators.

Methods

The Guideline, Barriers, Facilitators and Implementation strategies

Guideline development

The guideline was developed in 2000 using evidence-based guideline development principles.[14] Details are published elsewhere.[8] Recommendations (Table 1) were formu-

Table 1 The Guideline's Recommendations

1	In neonates and children with hypovolemia the first choice fluid for initial resuscitation is isotonic saline (Grade A)*
2	When large amounts of fluids are required (e.g. sepsis), it is possible to use a synthetic colloid because of its longer duration in the circulation (Grade C)
3	The initial fluid volume should be 10-20 ml/kg and repeated doses based on individual clinical response (Grade C)

* In adults there is Grade A evidence. This evidence is thought to be applicable to neonates and children.

lated by a national multidisciplinary committee comprising 26 members including the heads of all pediatric (n=8) and neonatal intensive care departments (n=10), general pediatricians (n=4), nurses (n=2) and other (n=2).

Questionnaire surveys

Questionnaires were developed to investigate pediatric practice of fluid solutions used for volume resuscitation in hypovolemia and to identify potential barriers and facilitators for guideline implementation. Before guideline development we surveyed *all* Dutch academic neonatal and pediatric intensive care departments (Figure 1). In 2004 we received a grant to start a nationwide implemention project for the guideline. To be able to distinguish on which groups we should focus our implementation strategies, and, to identify barriers and facilitators for implementation, we now involved a broader group of specialist treating children with hypovolemia from *all* academic and *all* general hospitals in the Netherlands (Figure 1). A randomly chosen pediatric specialist from each department was asked for fluid prescribing behavior on their departments. A covering letter stated that the results would be analyzed anonymously and non-responders were called after 6 weeks.

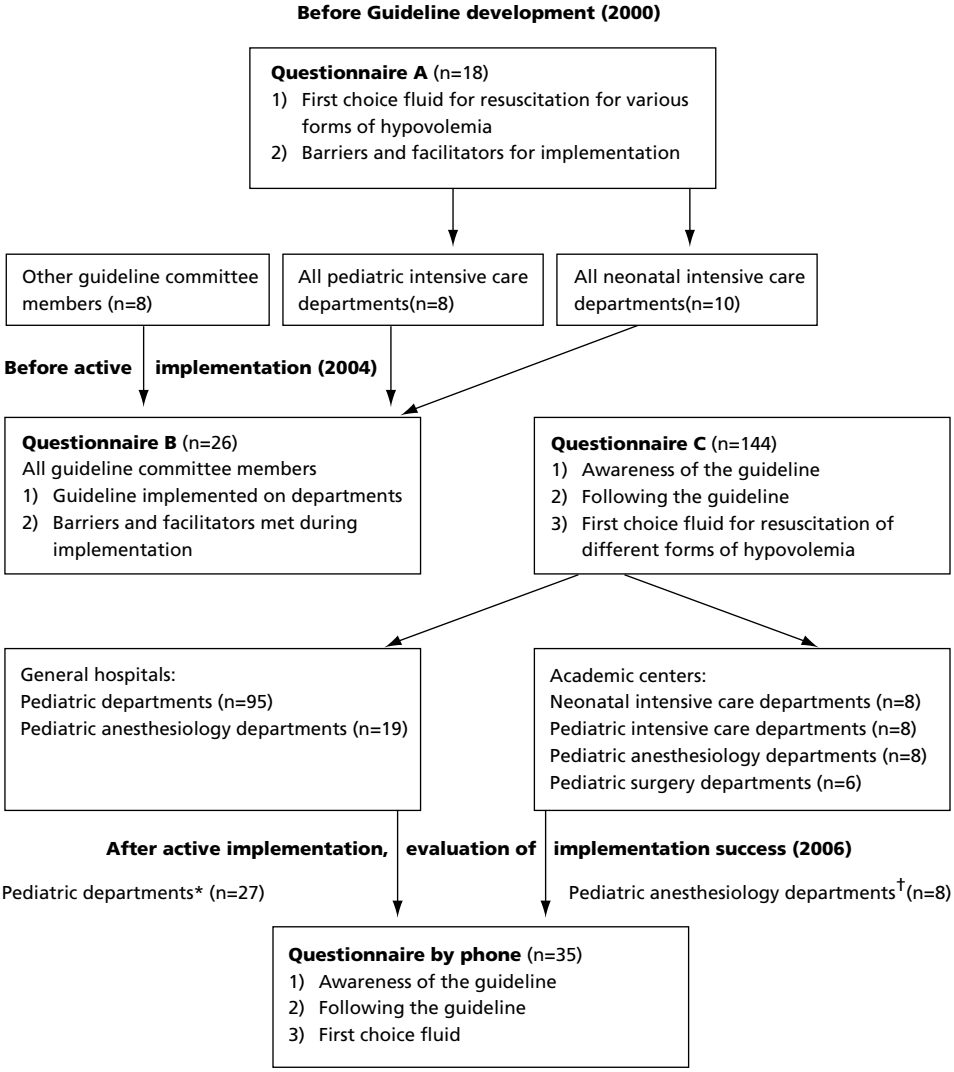
Evaluation of implementation success

The following quality-of-care indicators were developed: percentage of specialists aware of the guideline and percentage of specialists following the guideline. Quality-of-care indicators were measured using questionnaires and data from medical records. Successful implementation was defined as a 10% change in quality-of-care indicators after active implementation.

Reported behavior

After active implementation we evaluated implementation success in the group of specialists not always following the guideline's recommendations and/or those with a relatively high frequency of fluid use (Figure 1).

Figure 1 Flow diagram of surveys by questionnaire to different specialists before guideline development, after guideline development and after active implementation



* general hospital pediatric departments that had reported not to always follow the guideline's recommendations

† all academic departments of pediatric anesthesiology because they had the lowest use of crystalloids as first-choice volume therapy of all academic specialists, and because their frequency of fluid use is relatively high

Medical records

The initial plasma volume expander used was extracted from the medical records of 120 consecutive patients who were treated for hypovolemia in January and February 2006 on three academic neonatal and three academic pediatric intensive care units

(n=6): 1) one with the highest amount of albumin use before guideline development; 2) one with the highest amount of non-human colloids use, and, 3) one with moderate use of all plasma volume expanders compared with other academic centers. With a number of 20 medical records per academic hospital unit concerning 20 hypovolemic children in all of whom normal saline is prescribed as first-line treatment, one can conclude with a 95% CI of 84-100% that for all hypovolemic children on that department normal saline is used as first-line treatment.

Data-analysis

Statistical analysis included descriptive statistics and frequency analysis. All analyses were performed using SPSS 12.0 (SPSS, Chicago, USA).

Results

Questionnaire surveys

Response rate

The response rates to *Questionnaire A* and *B* were 95% (17/18) and 77% (20/26), respectively. The response rate to questionnaire C was 99% (143/144).

Barriers to and facilitators for guideline implementation and implementation strategies used

Table 2 shows barriers and facilitators to implementation reported during the whole process of guideline development and implementation. Table 3 shows all implementation strategies used, specifically targeting identified barriers.

Table 2 Reported implementation barriers and facilitators

I	**Barriers**
•	Uncertain applicability of evidence from adults to children
•	Fear of inducing hypernatremia in neonates when using normal saline
•	Existence of a local protocol
•	Ongoing recommendation to use albumin or synthetic colloids by international 'opinion leaders'
•	'Old habits and routines' that seem to work well
•	Unclear accessibility of the new guideline, lack of detailed knowledge about guideline's recommendations, reported 'unclear recommendations' and 'extensive guideline'
II	**Facilitators**
•	'Evidence-based' guideline, relevant to practice, with clear recommendations
•	Motivated members of multidisciplinary committee
•	Digital free accessibility of the guideline

Table 3 Implementation strategies

I At guideline development onset
- Personal interviews with academic specialists aimed at identification of potential barriers and facilitators for implementation
- Formation of multidisciplinary guideline development committee consisting of all stakeholders' representatives

II During guideline development
- Use of Evidence-based guideline development process
- Potential barriers to implementation addressed in guideline
- Involvement of all stakeholders in formulating the final guideline recommendations
- Local implementation by stakeholders on their departments and in regional hospitals

III After guideline development
- Endorsement of guideline by the Dutch Association of Pediatrics
- Educational visits of academic specialists
- Information stand at Dutch annual pediatric meeting; checking participants' recommendation knowledge offering a reward (plastic duck filled with bath salts)
- Proposed incentive: printed copy of full text guideline for those who fill out questionnaires
- All academic pharmacies received a printed version of the guideline
- Dissemination of the guideline's recommendation as a pocket-size plasticized card
- Digital free access to the full text guideline
- Publication of the guideline's recommendation in the Newsletter of the Dutch Association of Pediatrics and Dutch Journal of Pediatrics
- Interactive lessons about guideline recommendations for all Dutch pediatric residents

First choice fluid therapy before and after guideline development

See Figure 2a-d. Before guideline development in 2000 first choice fluid therapy on the NICU's for the different forms of shock was a crystalloid in 22-89%, albumin in 11-44% and a synthetic colloid in 0-11%. After guideline development in 2004 the reported use of these agents was 100%, 0%, and 0%, respectively. Before guideline development first choice fluid therapy on the PICU's for the different forms of shock was a crystalloid in 43-71%, albumin in 0-14% and a synthetic colloid in 29-43%. After guideline development the reported use of these agents was 88-100%, 0%, and 0-13%, respectively. After guideline *development*, but before *active implementation* reported first choice fluid therapy in general hospital pediatric departments for the different forms of shock was a crystalloid in 84-100%, albumin in 0-1% and a synthetic colloid in 0-2%; in academic pediatric surgical departments it was 33-100%, 0-17% and 0-17%, respectively and in academic pediatric anesthesiology departments it was 50-88%, 0% and 13-38%.

Local implementation and barriers and facilitators met by members of the multidisciplinary guideline development committee

Up to 95% of respondents of the guideline development committee reportedly implemented the guideline after 2001. Only 15% (3/20) reported barriers during their implementation consisting of 'a too extensive guideline text' and 'routines or old habits'.

Figure 2a Proportion first choice fluid in dehydration before (2000), after guideline development (2004) and after active implementation (2006) according to specialism

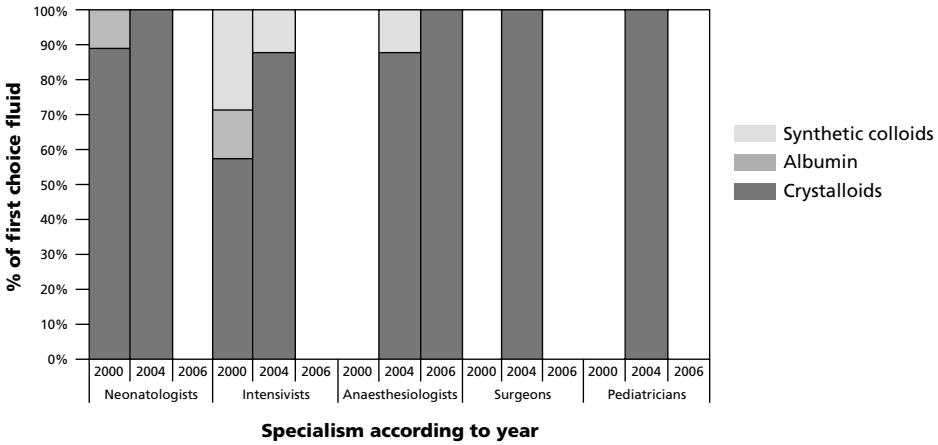
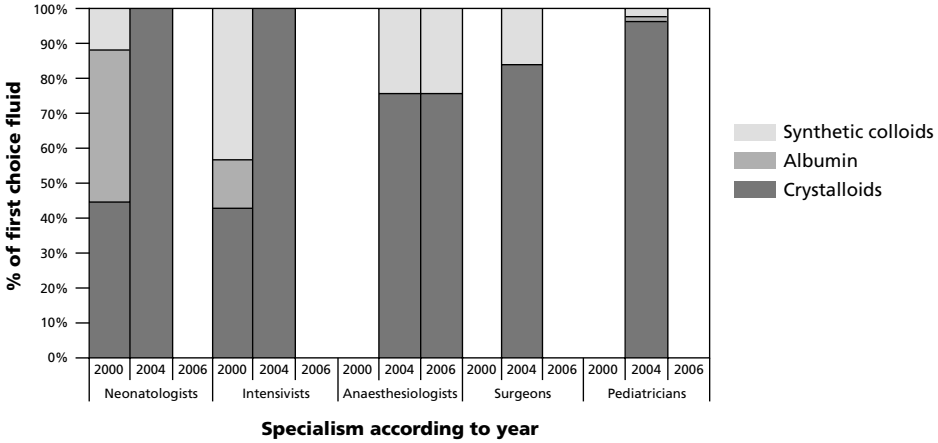


Figure 2b Proportion first choice fluid in sepsis before (2000), after guideline development (2004) and after active implementation (2006) according to specialism



Frequency of prescribing fluids for volume therapy

On average, academic pediatric specialists prescribe fluid therapy daily or several times a week. General hospital pediatricians prescribe fluid therapy monthly to once in every three months. General pediatric anesthesiologists prescribe initial fluid therapy once in three months.

Awareness of guideline by pediatric specialists not involved in the guideline development

Before active implementation 50% (72/144) of all practitioners, mainly general hospital pediatricians 66% (63/95), were aware of the guideline.

Figure 2c Proportion first choice fluid in trauma before (2000), after guideline development (2004) and after active implementation (2006) according to specialism

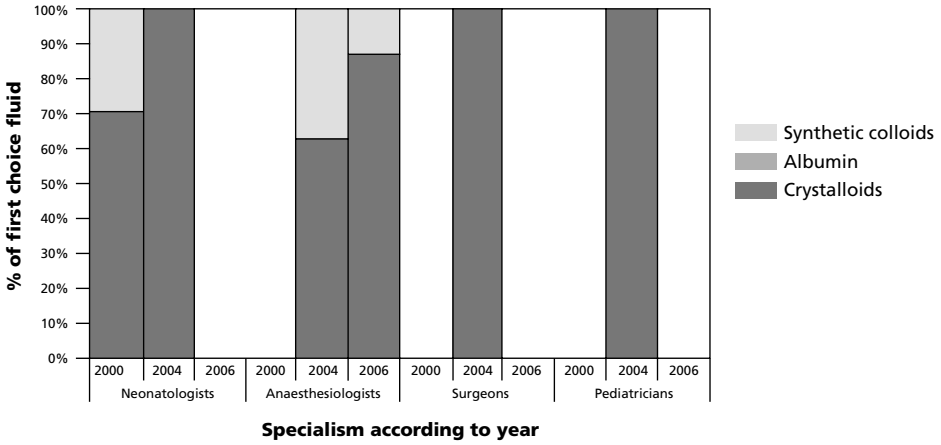
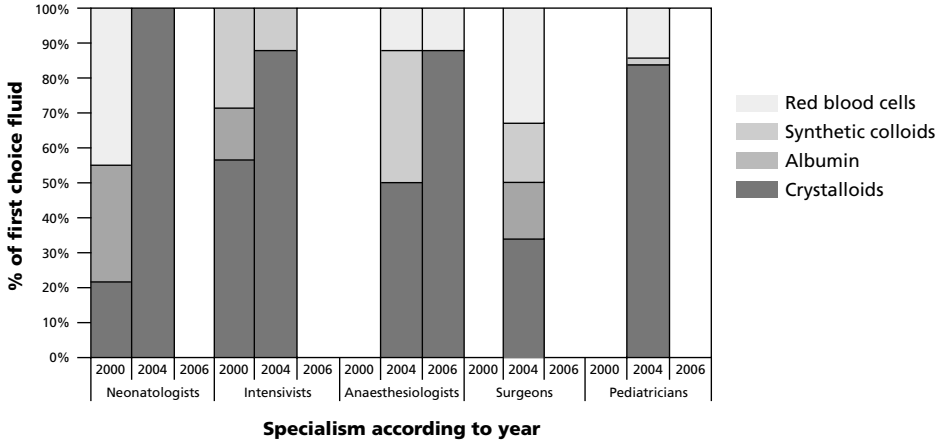


Figure 2d Proportion first choice fluid in hemorrhage before (2000), after guideline development (2004) and after active implementation (2006) according to specialism



Data were not obtained for all years for each different specialism (see Methods)

Academic specialists: NICU: n=10 in 2000, n=8 in 2004, neonatologists do not encounter trauma as a cause for hypovolemia in neonates; PICU: n=7 in 2000, n=8 in 2004; pediatric anesthesiology departments: n=8 in 2004 and 2006; pediatric surgery departments: n=6 in 2004. General hospital pediatricians: n=95 in 2004. Only a minority of specialists used other crystalloids (i.e. Ringer's solution and Ringer's lactate) instead of normal saline as first choice. Although normal saline was recommended in the guideline, for ease of reading and interpretation we report 'total crystalloid use'.

Pediatric anesthesiologists working in general hospitals are not reported separately, as they appeared to be infrequent users of fluid therapy for hypovolemia

Adherence to the guideline

Of academic specialists, 71% (10/14) always followed the recommendations and 29% (4/14) inconsistently. Reporting to *Questionnaire C,* 57% (36/63) of general pediatricians who were aware of the guideline claimed to 'always follow the recommendations' and 43% (27/63) 'inconsistently'.

Evaluation using indicators

Reported behavior

Before active implementation in 2004 only 13% (1/8) of academic pediatric anesthesiologist were familiar with the guideline compared to 50% (4/8) after active implementation in 2006. In this group, reported first choice fluid in 2006 was a crystalloid in 75-100%, albumin 0%, and synthetic colloids 0-25%. Before active implementation 43% (27/63) of general hospital pediatricians aware of the guideline did not always follow the guideline. After active implementation 89% (24/27) claimed to 'always' follow the recommendations. 11% (3/27) was not aware of the guideline but two of these three hospitals already used normal saline as first-choice treatment. For first choice fluid used before and after active implementation see Figure 2a-d.

Medical records

A total of 120 medical records from six departments were included. This showed that 109/120 (92%) of hypovolemic children received normal saline and that 5 out of 6 departments used normal saline in all 20 consecutive cases for initial treatment of hypovolemic shock. Only one pediatric intensive care department, which belonged to the academic center with the highest volume of non-human colloid use, did not always use normal saline.

Discussion

The publication and subsequent active implementation of a national practice guideline was followed by a change in prescribing practice by physicians treating children with hypovolemia in the Netherlands. Reported changes in care after guideline development and implementation are not that overwhelming.[16] A recent systematic review of 235 studies of guideline implementation strategies and their impact on practice revealed a median 10% change in care habits.[11]

Evidence and marketing in fluid prescribing behavior

We are the first to report the prescribing pattern of fluids for hypovolemia in a large number of specialists caring for sick children. Before guideline development we found variability in prescribing patterns: a considerable amount of neonatologists used albumin and pediatric intensivists predominantly used a synthetic colloid. Interestingly, the reported use of crystalloids and colloids differed for the different causes of shock. According to the available research evidence, there are no clear reasons for this, and so this is probably due to physician's beliefs and 'old habits'.

Several surveys have shown that, despite the publication of a number of systemat-

ic reviews [2-4;6;17-19] and the recent methodologically sound and large SAFE trial showing that colloids are not superior to crystalloids,[20] this evidence has not changed clinical practice: the majority of physicians still use costly colloid products.[7;21;22] One survey among pediatric specialists (i.e. pediatric anesthesiologists) showed that albumin continues to be the most frequently used volume expander in neonates, whereas alternative colloids, i.e. hetastarch and gelatins, are used in infants and older children.[7] Physicians reporting hetastarch use were more likely to report a visit from a drug detailer for hetastarch.[21] Several authors have shown that physicians may change their prescribing practice as a result of contact with drug detailers, irrespective of scientific evidence.[23-26] As we all know, powerful industries have impressive control over dissemination and implementation of information. Marketing strategies by the industry involves regular visits of stakeholders and identifying barriers. We feel we could learn from the industry's success of social marketing, but believe that clinicians should take the lead in implementing unbiased research.

Guideline development and implementation

In response to the uncertainty and variability in clinical practice regarding the optimal first choice fluid for volume resuscitation in children, we developed an evidence-based guideline endorsed by the Dutch Association of Pediatrics. A systematic literature review was performed and all stakeholders were involved in formulating the guideline recommendations after two consensus meetings. They were asked to implement the guideline on their departments and their regional local hospitals. Barriers identified were addressed in the guideline. After guideline development and before active implementation, all (100%) neonatologists reportedly used normal saline as a first choice fluid for resuscitation and 88% of pediatric intensivists; albumin use for resuscitation had declined to zero. Our survey in a broader group of pediatric specialists showed that before active implementation most of the pediatric specialist's management behavior was already in keeping with the guideline's recommendations. We presume that stakeholders who were involved in the developmental process have been of great importance in disseminating the recommendations in their region. Indeed, 95% of the guideline development committee members reportedly implemented the guideline in their own departments, including all national pediatric and neonatal intensive care units.

There are well-defined steps that should be taken to promote the clinical uptake of research findings.[27;28] Once there is uncertainty about the most appropriate treatment a guideline should be developed according to evidence-based principles. High quality guidelines are based on evidence as well as broad consensus of opinions, which facilitates the acceptance and effective use in the target group.[29] The following guideline characteristics contribute to their use in practice: inclusion of specific recommendations, sufficient supporting evidence, a clear structure, an attractive lay out and short

summaries.[30;31] Guideline endorsement by a physician's own specialty organization has been associated with improved physician's confidence in a guideline.[32-34] Pediatricians rely on their close colleagues, particularly local sub-specialists, before they change their practice habits.[35]

Surveys of pediatrician's attitudes towards guidelines in general report barriers to adherence such as lack of awareness, lack of agreement with specific recommendations, or lack of agreement with the concept of guidelines in general.[36] The same survey showed that recent graduates reported change of behavior more often than more seasoned practitioners.[36] Identification of these barriers can foster opportunities to improve physician adherence and thus reduce variability in practice, improve patient care or reduce costs. Physician participation in guideline development has been shown to be useful in addressing barriers owing to lack of agreement.[37]

Implementation strategies and Study design

Several systematic reviews including hundreds of studies have taught us that there is no 'magic bullet' for implementation success.[10;11;38-40] Based on systematic reviews published by the Cochrane Effective Practice and Organization of Care Group (EPOC) implementation strategies can be categorized into three groups showing consistent, variable, or little or no effectiveness. Those interventions that consistently have shown effectiveness include interactive educational meetings, educational outreach visits, reminders, and multifaceted interventions (defined as a minimum of two combined interventions). Interventions that have shown variable effectiveness include audit and feedback, the use of local opinion leaders, local consensus processes, and patient-mediated interventions. Interventions that have consistently showing little or no effect are didactic educational meetings (lecture-format) and educational materials.[10;11;38]

We decided not to perform a randomized trial comparing various different implementation strategies because we believe the results will not be applicable to other future guidelines. In a time where healthcare budgets are limited, we should carefully consider the most appropriate design of a study. There is a strong call for multiple arm cluster randomized controlled trials to investigate the (cost) effectiveness of different implementation strategies.[41;42] An important advantage of this design is high internal validity, disadvantages are the fact that they are expensive, time consuming, and, generally, will have low external validity and applicability in the case of comparing implementation strategies for guidelines. First of all, the success of implementation is very much dependent on the guideline development process, e.g. method of guideline development, strength of the underlying evidence, belief in the members of the guideline development group, etc. We have shown that by developing a robust guideline involving all stakeholders, even without using active implementation strategies we achieved a considerable change in behavior. Secondly, the effect of different implementation strategies like educational meetings, outreach visits and the use of local

opinion leaders, will all highly depend on the commitment and skills of the persons executing these strategies. Their effectiveness and importance cannot be overstated, but, unfortunately cannot be quantitated nor isolated as uncontrolled independent variables. As all these issues will be guideline specific, we decided it was more cost-effective and efficient to continuously develop and adjust implementation strategies according to identified barriers (Table 1 and 2).

Explicit cost-containment in guidelines

Cost-impact assessments are potentially problematic since they may be perceived to take the focus away from improving patient care, potentially discrediting the guideline medium. Our first aim was to improve patient care. Different treatments proved to be equally effective, so we chose the cheapest. All academic pharmacists received the guideline as an implementation strategy. They could have played an important role by restricting albumin use for other indications than hypovolemia. A formal cost-minimization analysis was not possible, as we could not obtain valid pharmaceutical data. What we do know is that for every Liter normal saline used instead of albumin, 138 Euro (150 US$) is saved and for every Liter normal saline used instead of a synthetic colloid (e.g. HES), 23 Euro (25 US$) is saved.

Guidelines focusing on interventions with equal effectiveness but different financial costs, should invite health insurers, health policy makers and pharmacists for their multidisciplinary committee. In order to perform a cost-minimization analysis, indicators should be developed *before* guideline development and checked for availability and validity.

Evaluation of implementation success by indicators

No patient oriented outcomes could be used, because in this case the evidence showed no difference between treatments. Therefore, making a difference in costs was the main objective. We tried to collect pharmaceutical data on the annual volume of different fluid solutions used in neonates and children from hospital registries. This appeared to be impossible as fluid therapy is not registered on indication or on number of patients and most often not separately for different departments.

According to the questionnaires, active implementation was successful. As self-reported measures potentially reveal an idealized version of actual behavior, we developed a more objective indicator: medical records. We selected academic hospitals with the highest colloid use before guideline implementation as they are most likely to report idealized behavior, i.e. saline use, after guideline implementation. It is not very likely that hospitals already reporting saline use before guideline implementation use colloids after guideline implementation. Data from medical records were in keeping with answers from the questionnaires: 92% of hypovolemic children actually received normal saline after active implementation.

Limitations

A disadvantage of our uncontrolled design is the fact that secular trends or sudden changes due to external factors make it difficult to attribute observed changes solely to the intervention. Changes in evidence (e.g. publication of the SAFE Study) could have contributed to the change in fluid prescribing behavior. However, as we pointed out earlier, several surveys have shown that such evidence did not change local clinical practice. In addition, colloid use after guideline development and implementation was quite different compared to other recent surveys on fluid prescribing behavior. Therefore, we believe that guideline development and implementation were important factors in the change of fluid prescribing behavior.

Another possible limitation is the fact that we evaluated post-active implementation (2006) success in a limited group of specialists. Because most of the reported physician's prescribing behavior before active implementation was already in keeping with the guideline, we decided it was not cost-effective to extensively evaluate fluid prescribing behavior in all specialists after active implementation. Again, it seems very unlikely that physicians already using normal saline as a first choice fluid before active implementation, will use a colloid as a first choice fluid after guideline implementation. To be sure that reported behavior reflects actual behavior, we obtained data from medical records in 2006 showing that indeed 92% of children in university hospitals received normal saline as a resuscitation fluid. We therefore believe we can rely on the answers obtained from the questionnaires.

Conclusions and recommendations

Evidence-based guidelines have the power to change clinical practice of pediatric specialist, even if it is only for cost-containment. To successfully implement clinical guidelines and reduce the huge cost of active implementation, any guideline development should consider implementation right from the beginning. This implies involvement of all stakeholders, addressing potential barriers, credibility, transparent structure in the information, i.e. systematic review of the literature and clearly presenting all (other) considerations that influenced the final recommendation. Finally an attractive, accessible format with brief summaries should be made available. Implementation strategies should be multifaceted, i.e. tailored to suit local circumstances and taking into account any particular barrier identified. Success or failure of implementation should be measured by guideline-specific indicators. The health care sector should be made aware that in the process of improving quality of care appropriate and easily available measurements are urgently needed.

Acknowledgements

The authors would like to thank Pim Langendijk, pharmacist, for developing a spreadsheet form of different plasma volume expanders and given support by collecting data of academic pharmacists, and Dr Joke Kok, MD PhD and Dr Jako Burgers, MD PhD for their helpful comments on an earlier version of the manuscript.

This project was funded by the Netherlands' Order of Medical Specialists.

Reference List

1 Offringa M. Excess mortality after human albumin administration in critically ill patients. Clinical and pathophysiological evidence suggests albumin is harmful. BMJ 1998; 317(7153):223-224.

2 Human albumin administration in critically ill patients: systematic review of randomised controlled trials. Cochrane Injuries Group Albumin Reviewers. BMJ 1998; 317 (7153):235-240.

3 Schierhout G, Roberts I. Fluid resuscitation with colloid or crystalloid solutions in critically ill patients: a systematic review of randomised trials. BMJ 1998; 316(7136):961-964.

4 Alderson P, Schierhout G, Roberts I, Bunn F. Colloids versus crystalloids for fluid resuscitation in critically ill patients. Cochrane Database Syst Rev 2000;(2):CD000567.

5 Bunn F, Alderson P, Hawkins V. Colloid solutions for fluid resuscitation. Cochrane Database Syst Rev 2001;(2):CD001319.

6 Choi PT, Yip G, Quinonez LG, Cook DJ. Crystalloids vs. colloids in fluid resuscitation: a systematic review. Crit Care Med 1999; 27(1):200-210.

7 Soderlind M, Salvignol G, Izard P, Lonnqvist PA. Use of albumin, blood transfusion and intraoperative glucose by APA and ADARPEF members: a postal survey. Paediatr Anaesth 2001; 11(6):685-689.

8 Boluyt N, Bollen CW, Bos AP, Kok JH, Offringa M. Fluid resuscitation in neonatal and pediatric hypovolemic shock: a Dutch Pediatric Society evidence-based clinical practice guideline. Intensive Care Med 2006; 32(7):995-1003.

9 Grol R, Grimshaw J. From best evidence to best practice: effective implementation of change in patients' care. Lancet 2003; 362(9391):1225-1230.

10 Grimshaw JM, Shirran L, Thomas R, Mowatt G, Fraser C, Bero L et al. Changing provider behavior: an overview of systematic reviews of interventions. Med Care 2001; 39(8 Suppl 2):II2-45.

11 Grimshaw JM, Thomas RE, MacLennan G, Fraser C, Ramsay CR, Vale L et al. Effectiveness and efficiency of guideline dissemination and implementation strategies. Health Technol Assess 2004; 8(6):iii-72.

12 Kuhlthau K, Ferris TG, Iezzoni LI. Risk adjustment for pediatric quality indicators. Pediatrics 2004; 113(1 Pt 2):210-216.

13 Lawrence M, Olesen F. Indicators of quality in health care. Eur J Gen Pract 1997; 3:103-108.

14 National Health and Medical Research Council. A Guide to the Development, Implementation and Evaluation of Clinical Practice Guidelines. Canberra: NHMRC, 1999.

15 Tabbers MM, Boluyt N, Offringa M. Fluid resuscitation in neonatal and pediatric hypovolemic shock: an evidence-based guideline. Dutch Journal of Pediatrics 2004; 6(72):242-243.

16 Grol R. Successes and failures in the implementation of evidence-based guidelines for clinical practice. Med Care 2001; 39(8 Suppl 2):II46-II54.
17 Alderson P, Bunn F, Lefebvre C, Li WP, Li L, Roberts I et al. Human albumin solution for resuscitation and volume expansion in critically ill patients. Cochrane Database Syst Rev 2002;(1):CD001208.
18 Bisonni RS, Holtgrave DR, Lawler F, Marley DS. Colloids versus crystalloids in fluid resuscitation: an analysis of randomized controlled trials. J Fam Pract 1991; 32(4):387-390.
19 Velanovich V. Crystalloid versus colloid fluid resuscitation: a meta-analysis of mortality. Surgery 1989; 105(1):65-71.
20 Finfer S, Bellomo R, Boyce N, French J, Myburgh J, Norton R. A comparison of albumin and saline for fluid resuscitation in the intensive care unit. N Engl J Med 2004; 350 (22):2247-2256.
21 Miletin MS, Stewart TE, Norton PG. Influences on physicians' choices of intravenous colloids. Intensive Care Med 2002; 28(7):917-924.
22 Schortgen F, Deye N, Brochard L. Preferred plasma volume expanders for critically ill patients: results of an international survey. Intensive Care Med 2004; 30(12):2222-2229.
23 Chren MM, Landefeld CS. Physicians' behavior and their interactions with drug companies. A controlled study of physicians who requested additions to a hospital drug formulary. JAMA 1994; 271(9):684-689.
24 Lurie N, Rich EC, Simpson DE, Meyer J, Schiedermayer DL, Goodman JL et al. Pharmaceutical representatives in academic medical centers: interaction with faculty and housestaff. J Gen Intern Med 1990; 5(3):240-243.
25 Orlowski JP, Wateska L. The effects of pharmaceutical firm enticements on physician prescribing patterns. There's no such thing as a free lunch. Chest 1992; 102(1):270-273.
26 Peay MY, Peay ER. The role of commercial sources in the adoption of a new drug. Soc Sci Med 1988; 26(12):1183-1189.
27 Grol R, Grimshaw J. Evidence-based implementation of evidence-based medicine. Jt Comm J Qual Improv 1999; 25(10):503-513.
28 Haines A, Donald A. Making better use of research findings. BMJ 1998; 317(7150):72-75.
29 Burgers JS. Guideline quality and guideline content: are they related? Clin Chem 2006; 52(1):3-4.
30 Wollersheim H, Burgers J, Grol R. Clinical guidelines to improve patient care. Neth J Med 2005; 63(6):188-192.
31 Grol R, Dalhuijsen J, Thomas S, Veld C, Rutten G, Mokkink H. Attributes of clinical guidelines that influence use of guidelines in general practice: observational study. BMJ 1998; 317(7162):858-861.
32 Hayward RS, Guyatt GH, Moore KA, McKibbon KA, Carter AO. Canadian physicians' attitudes about and preferences regarding clinical practice guidelines. CMAJ 1997; 156(12):1715-1723.

33 James PA, Cowan TM, Graham RP, Majeroni BA. Family physicians' attitudes about and use of clinical practice guidelines. J Fam Pract 1997; 45(4):341-347.

34 Pope C, Mays N. Reaching the parts other methods cannot reach: an introduction to qualitative methods in health and health services research. BMJ 1995; 311(6996):42-45.

35 Weiss R, Charney E, Baumgardner RA, German PS, Mellits ED, Skinner EA et al. Changing patient management: what influences the practicing pediatrician? Pediatrics 1990; 85(5):791-795.

36 Christakis DA, Rivara FP. Pediatricians' awareness of and attitudes about four clinical practice guidelines. Pediatrics 1998; 101(5):825-830.

37 Lagoe RJ, Aspling DL. Enlisting physician support for practice guidelines in hospitals. Health Care Manage Rev 1996; 21(4):61-67.

38 Bero LA, Grilli R, Grimshaw JM, Harvey E, Oxman AD, Thomson MA. Closing the gap between research and practice: an overview of systematic reviews of interventions to promote the implementation of research findings. The Cochrane Effective Practice and Organization of Care Review Group. BMJ 1998; 317(7156):465-468.

39 Grimshaw J, Eccles M, Thomas R, MacLennan G, Ramsay C, Fraser C et al. Toward evidence-based quality improvement. Evidence (and its limitations) of the effectiveness of guideline dissemination and implementation strategies 1966-1998. J Gen Intern Med 2006; 21 Suppl 2:S14-20.:S14-S20.

40 Oxman AD, Thomson MA, Davis DA, Haynes RB. No magic bullets: a systematic review of 102 trials of interventions to improve professional practice. CMAJ 1995; 153(10): 1423-1431.

41 Eccles M, Grimshaw J, Campbell M, Ramsay C. Research designs for studies evaluating the effectiveness of change and improvement strategies. Qual Saf Health Care 2003; 12(1):47-52.

42 Grimshaw J, Campbell M, Eccles M, Steen N. Experimental and quasi-experimental designs for evaluating guideline implementation strategies. Fam Pract 2000; 17 Suppl 1:S11-6.:S11-S16.

43 Brand C, Landgren F, Hutchinson A, Jones C, Macgregor L, Campbell D. Clinical practice guidelines: barriers to durability after effective early implementation. Intern Med J 2005;35(3):162-169.

CHAPTER 7

Quality of evidence-based pediatric guidelines

Nicole Boluyt, Carsten R. Lincke, Martin Offringa

Pediatrics 2005 May;115(5):1378-9

ABSTRACT

Objective To identify evidence-based pediatric guidelines and to assess their quality.

Methods We searched MEDLINE, EMBASE, and relevant Web sites of guideline development programs and national pediatric societies to identify evidence-based pediatric guidelines. A list with titles of identified guidelines was sent to 51 leading pediatricians in the Netherlands, who were asked to select the 5 most urgent topics for guideline development. Three pediatrician reviewers appraised the available guidelines on the 10 most frequently mentioned topics with the Appraisal of Guidelines for Research and Evaluation (AGREE) instrument.

Results A total of 215 evidence-based pediatric guidelines were identified; of these, 17 guidelines on the 10 most frequently mentioned topics were appraised. The AGREE instrument rates guidelines among 6 domains. For the scope and purpose domain, the mean score was 84% of the maximum mark. For stakeholder involvement, the mean score was 42%, with 12 guidelines (71%) scoring <50%. For rigor of development, the mean score was 54%, with 5 (29%) guidelines scoring <50%. For clarity and presentation, the mean score was 78%, with 4 (24%) guidelines scoring <50%. For applicability and editorial independence, performance was poor, with mean scores of 19% and 40%, respectively. Low scores were partly attributable to poor reporting. After considering all domain scores, the reviewers recommended 14 of 17 guidelines (82%) to be used in local practice.

Conclusions The current volume of pediatric guidelines categorized as evidence based in popular databases is large. Overall, these guidelines scored well, compared with other studies on guideline quality in fields outside pediatrics, when assessed for quality with the AGREE instrument. This holds especially for guidelines published or endorsed by the American Academy of Pediatrics or registered in the National Guideline Clearinghouse.

Introduction

Clinical practice guidelines are systematically developed statements to assist practitioners and patients in making decisions about appropriate health care in specific clinical circumstances.[1] For many health care conditions, a gap exists between what medical science has shown to be effective practice and what is actually done.[2] The primary goal of practice guidelines in pediatrics is to improve the health of infants and children by ensuring they receive up-to-date, evidence-based care. Practice guidelines represent one of the various tools that can be used to improve the quality of care.[3] Several studies have shown that adherence to evidence-based guidelines leads to improvement in the quality of care provided.[4-10]

In recent decades, the number of available clinical practice guidelines has grown enormously. It is estimated that ~ 2500 guidelines are already in existence. This recent increase in the production of clinical practice guidelines has been accompanied by growing concern about the variations in guideline recommendations and quality. In fact, several studies suggest that many existing guidelines are of poor quality.[11-15] As the number of published guidelines proliferates there have been calls for the establishment of internationally recognized standards to improve the development and reporting of clinical guidelines. For this purpose, an international group of researchers from 13 countries, the Appraisal of Guidelines, Research, and Evaluation (AGREE) Collaboration, has developed and validated a generic instrument that can be used to assess the quality of clinical guidelines.[16]

The number of published pediatric guidelines available to pediatricians is also increasing rapidly. However, their quality has not been assessed systematically. We set out to measure the current volume of potentially high-quality, published, pediatric guidelines, to assess their quality with the AGREE instrument, and to see whether we could adjust them for local use (to avoid potential duplication of effort within our Dutch national guideline development program). We limited our search to evidence-based guidelines, to enhance the yield of high-quality guidelines. We asked the following 2 questions: What evidence-based pediatric guidelines exist currently? What is the quality of these guidelines?

Methods

Literature Search

We searched MEDLINE (1966 to January 2004) and EMBASE (1988 to January 2004) with the following terms: evidence-based medicine (Medical Subject Heading) or evidence based.tw (text word) or evidence-based.tw and practice guidelines.pt (publication type), with the limits 'child' and 'English language'. Furthermore, we looked at relevant Web sites of agencies known to produce and/or endorse evidence-based guidelines, i.e.

American Academy of Pediatrics (AAP) and Scottish Intercollegiate Guidelines (SIGN), and Web sites of databases known to register evidence-based guidelines, i.e. the Agency for Health Care Policy and Research National Guideline Clearinghouse (NCG) and the Canadian Medical Association (CMA) Clinical Practice Guidelines Infobase.

Inclusion of guidelines

Included were practice guidelines that concerned the management of diseases among children and neonates. Not all agencies that produce guidelines focus solely on the pediatric age group and, in order to be included in our sample, guidelines were required to make specific recommendations for children and/or neonates. Excluded were guidelines on prevention, screening, prenatal diagnostics, psychiatry, surgery, general practice, and very rare disorders. One of our aims was to appraise critically a sample of the retrieved guidelines, to see whether we could adopt them for use in the Netherlands. We sent the list with evidence-based guidelines that had been identified through the literature search to all directors of pediatric training programs in both academic teaching hospitals (n=10) and affiliated teaching hospitals in community practice settings (n=22) and to the chairpersons of all the subspecialty working groups of the Dutch Board of Pediatricians in the Netherlands (n=19), and we asked them to select the 5 most urgent topics for national pediatric guideline development.

Appraisal of guidelines with the AGREE instrument

We appraised the evidence-based guidelines published up to 2002 that were identified through the literature search for the 10 topics mentioned most frequently by 51 Dutch leaders in pediatrics, with 3 separate reviewers (N.B., C.R.L., and M.O.) using the AGREE instrument to evaluate the scientific quality of the selected guidelines. We collected all documentation related to each guideline available in the public domain. Before we started to appraise the guidelines, we appraised a practice set to reach consensus about the interpretation of all items. The AGREE instrument is an international, methodologically rigorously developed, validated instrument.[16] It contains 6 domains, with a total of 23 items, and allows for the assessment of several components that are integral to guideline development, as follows: (1) scope and purpose (3 items), (2) stakeholder involvement (4 items), (3) rigor of development (7 items), (4) clarity and presentation (4 items), (5) applicability (3 items), and (6) editorial independence (2 items) (Appendix 1). The score for each domain is obtained by summing all of the scores of the individual items in a domain and then standardizing as follows: (obtained score – minimal possible score)/(maximal possible score – minimal possible score). The maximal score for each domain would be the number of questions multiplied by the number of reviewers times 4 (i.e. the score for 'strongly agree'). The minimal possible score for a domain would be the number of questions multiplied by the number of reviewers times 1 (i.e. the score for 'strongly disagree').

The final item of the AGREE instrument involves a recommendation regarding the use of the guideline in practice, as 'strongly recommended', 'recommended with provisos or alterations', 'would not recommend', or 'unsure'. For ease of interpretation, we considered 'strongly recommended' and 'recommended with provisos or alterations' as responses indicating 'recommended' and 'would not recommend' or 'unsure' as responses indicating 'not recommended'.

Cohen's κ, a widely used measure of agreement, was considered. For this purpose, the AGREE response categories were dichotomized into strongly agree/agree versus strongly disagree/disagree, because we judged an analysis of agreement at this level to be sufficient. However, substantial imbalances in the distribution of the tables' marginal totals were present, making comparison and interpretation of the κ values according to Cohen's criteria inappropriate.[17;18] Intraclass correlation coëfficients (ICCs) were considered to assess the inter-rater reliability within each domain. Similarly to κ values, ICCs are dependent on the variance of the item scores and thus the resulting domain scores. Because of lack of variance of the obtained item scores for some domains in this study, resulting ICCs would be low and could therefore be misinterpreted as lack of agreement. For that reason, we report only the observed proportion of overall agreement among the reviewers for each of the 23 items of the AGREE instrument.

Results

Literature Search

A total of 215 guidelines were identified by the search process after application of inclusion and exclusion criteria (Appendix 2). We identified 51 relevant guidelines in MEDLINE and 15 in EMBASE. In the NGC, we identified 119 guidelines. The AAP had 26 guidelines (endorsed guidelines included) available that fulfilled our inclusion and exclusion criteria. The SIGN had 3 guidelines available that fulfilled our inclusion and exclusion criteria. On the Web site of the CMA Clinical Practice Guidelines Infobase, we identified 34 guidelines. Some of the guidelines were found in several databases. All the AAP and SIGN guidelines were also registered in the NGC. Seven guidelines were found in both the NGC and MEDLINE. Three guidelines were found in both MEDLINE and EMBASE.

Priority list

The response to our request to colleagues to provide a list of 5 topics that needed a Dutch national guideline was 47% (24 of 51 individuals) after 2 mailings. The 10 most frequently mentioned topics and clinical problems are listed in Table 1.

Table 1 Prioritized topics for guideline development

	Topics
1	Constipation
2	Urinary tract infections
3	Head injury among children <2 y old
4	Head injury among children >2 y old
5	Diabetic ketoacidosis
6	Sedation for procedures
7	Antiemetics for patients receiving chemotherapy
8	Vesicoureteral reflux
9	Fever without source among children <2 mo old
10	Fever without source among children >2 mo old

Appraisal of guidelines

We appraised 17 evidence-based guidelines regarding the 10 most frequently mentioned topics. One guideline on procedural sedation did not address the prioritized topic and therefore was not appraised. The quality of the guidelines is indicated by their scores in Table 2. Table 2 lists 19 guidelines published in 17 documents; 2 guidelines (1 in MEDLINE and 1 in EMBASE) were appraised for fever among children <2 months of age and children >2 months of age. We now describe the appraisal results according to AGREE domain.

Scope and Purpose

The score for this domain represents the degree to which the overall objectives of the guideline, the clinical questions covered, and the patients to whom the guideline was meant to apply were described specifically. Overall, the mean score was 84% (range: 59 -100%).

Stakeholder Involvement

This domain evaluates the degree to which the guideline represents the views of its intended users. Included are questions regarding the composition of the guideline development group (specifically, whether individuals from all relevant professional groups were represented), whether patients' experiences and expectations influenced the development of the guideline, whether the target users of the guideline were well defined, and whether the guideline was pilot-tested among end-users. Overall, the mean score for this domain was 42% (range: 17-69%), with 12 guidelines (71%) scoring <50%. Thirteen guidelines (76%) included individuals from all relevant professional groups in the guideline development stage, but none involved patients in the development or was pilot-tested among end-users.

Table 2 AGREE domain scores (%) for selected evidence-based guidelines

Guideline	Scope and Purpose	Stakeholder Involvement	Rigor of Development	Clarity of Presentation	Appli-cability	Editorial Independence	Overall Recom-mendation
CONSTIPATION							
Idiopathic constipation and soiling in children (NGC)[21]	89	42	52	81	0	33	Recommended
Constipation in infants and children: evaluation and treatment (NGC, AAP endorsed)[22]	81	47	75	86	0	56	Recommended
URINARY TRACT INFECTIONS							
Diagnosis, treatment, and evaluation of the initial urinary tract infection in febrile infants and young children (AAP, NGC)[23]	100	58	75	86	52	44	Recommended
Evidence-based clinical practice guideline for patients ≤6 y old with a first-time acute urinary tract infection (NGC)[24]	81	44	52	86	15	28	Recommended
HEAD INJURY AMONG CHILDREN < 2 Y OLD							
Evaluation and management of children <2 y old with apparently minor head trauma (MEDLINE)[25]	96	39	57	92	19	33	Recommended
HEAD INJURY AMONG CHILDREN > 2 Y OLD							
Management of minor closed head injury in children (AAP)[26]	96	53	71	86	37	50	Recommended
Early management of patients with a head injury (NGC, SIGN)[27]	81	39	67	97	59	78	Recommended
Management of children with head trauma (CMA)[28]	59	19	2	36	0	44	Would not Recommend
DIABETIC KETOACIDOSIS							
Provincial protocol for the treatment of diabetic ketoacidosis in children (CMA)[29]	70	36	10	89	7	50	Recommended
ANTIEMETICS FOR PATIENTS RECEIVING CHEMOTHERAPY							
Recommendations for the use of antiemetics: evidence-based, clinical practice guidelines (NGC)[30]	74	50	81	81	11	67	Recommended

Guideline	Scope and Purpose	Stakeholder Involvement	Rigor of Development	Clarity of Presentation	Appli-cability	Editorial Independence	Overall Recom-mendation
ASHP therapeutic guidelines on the pharmacologicmanagement of nausea and vomiting in adult and pediatric patients receiving chemotherapy or radiation therapy or undergoing surgery (NGC)[31]	93	44	79	81	30	50	Recommended
VESICOURETERAL REFLUX							
Management of primary vesicoureteral reflux in children (NGC, AAP endorsed)[32]	93	69	84	89	30	56	Recommended
FEVER WITHOUT SOURCE AMONG CHILDREN < 2 MO OLD							
Practice guideline for the management of infants and children 0-36 mo old with fever without source (MEDLINE)[33]	93	31	59	94	22	33	Recommended
Evidence-based clinical protocol guideline for fever of uncertain source in infants ≤60 days old (NGC)[34]	85	53	62	89	15	17	Recommended
Evaluation of the infant with fever without source: an evidence-based approach (EMBASE)[35]	81	17	40	28	15	17	Would not Recommend
FEVER WITHOUT SOURCE AMONG CHILDREN > 2 MO OLD							
Practice guideline for the management of infants and children 0-36 mo old with fever without source (MEDLINE)[33]	93	31	59	94	22	33	Recommended
Evidence-based clinical practice guideline of fever of uncertain source: outpatient evaluation and management for children 2-36 mo old (NGC)[36]	85	47	49	89	7	17	Recommended
Evaluation of the infant with fever without source: an evidence-based approach (EMBASE)[35]	81	17	40	28	15	17	Would not Recommend
Management of the febrile 1-to 36-mo-old child with no focus of infection (CMA)[37]	67	19	11	31	0	17	Would not Recommend
Means	84	42	54	78	19	40	

NGC: www.guideline.gov; AAP: www.aap.org/policy/paramtoc.html; SIGN: www.sign.ac.uk; CMA: mdm.ca/cpgsnew/cpgs/index.asp

Rigor of Development

This domain evaluates specifically whether systematic methods were used to search for evidence, whether the criteria for selecting the evidence and the methods used to formulate the recommendations were clearly described, whether there was an explicit link between the recommendations and the supporting evidence, whether health benefits, side effects, and risks were considered when formulating the recommendations, whether the guideline was externally reviewed by experts prior to publication, and whether a procedure for updating the guideline was provided. Overall, the mean score for this domain was 54% (range: 2-84%), with 5 guidelines (29%) scoring <50%. Specifically, 14 guidelines (82%) described systematic methods for searching and selecting the evidence, 9 guidelines (53%) considered health benefits, side effects, and risks when formulating the recommendations, 11 guidelines (65%) described the methods used to formulate the recommendations, 14 guidelines (82%) indicated an explicit link between the supporting evidence and the recommendations, and 11 guidelines (65%) were reviewed externally before publication. Only 3 guidelines (18%) described a procedure for updating the guideline.

Clarity and Presentation

This domain describes the clarity of the guidelines. Specifically, it describes whether the recommendations were specific and unambiguous, whether the different management options were presented clearly, whether key recommendations were easily identifiable, and whether the guideline was supported with tools for application. Overall, the mean score for this domain was 78% (range: 28-97%). Four guidelines (24%) scored <50% for this domain.

Applicability

This domain evaluates the likely organizational, behavioral, and cost implications of applying the guideline. In addition, review criteria should be developed that link guideline use to audits and other quality improvement initiatives should be developed. The score on this domain was the lowest of all, with a mean score of 19% (range: 0-59%). Only 2 guidelines (12%) scored ≥50%. One guideline provided review criteria for monitoring purposes, and 3 discussed potential organizational barriers. Only 1 guideline discussed cost implications.

Editorial Independence

This domain addresses conflict of interest, specifically whether the guideline was editorially independent from the funding body and whether potential conflicts of interest were reported for the members of the guideline development group. The mean score in this domain was 40% (range: 17-78%). Seven guidelines (41%) scored >50%. In 15 guidelines (88%), potential conflicts of interest on the part of guideline developers were not recorded.

Agreement among reviewers

Of the 17 guidelines appraised, the reviewers agreed about the subjective judgment in 88% (15 guidelines) and came to a consensus with respect to an overall recommendation for each guideline. In total, we recommended 14 of 17 guidelines (82%). Table 3 summarizes the observed simple agreement among reviewers for the 23 items of the AGREE instrument. Observed agreement among reviewers was 41 to 60% for 3 items, 61 to 80% for 12 items, and >80% for 8 items.

Table 3 Agreement among 3 reviewers for AGREE instrument items

Agreement, %	No. of Items
0	0
0-20	0
21-40	0
41-60	3
61-80	12
>80	8

Discussion

Practice guidelines have become an increasingly popular tool for synthesis of clinical research that may be used to change clinical practice and to improve quality in health care. The quantitative growth in the number of guidelines available in different specialties is, however, a source of concern since there is evidence that the quality of these guidelines is generally poor.

We used the AGREE instrument to assess the quality of current pediatric guidelines. Mean domain scores in this survey were 84% for scope and purpose, 42% for stakeholder involvement, 54% for rigor of development, 78% for clarity of presentation, 19% for applicability and 40% for editorial independence. Compared with other studies of the quality of guidelines assessed with the AGREE instrument in fields outside pediatrics, these scores are fairly good. For example, the international validation survey of the AGREE instrument (http://www.openclinical.org/prj_agree.html) showed mean domain scores of 69% for scope and purpose, 36% for stakeholder involvement, 41% for rigor of development, 66% for clarity of presentation, 37% for applicability, and 30% for editorial independence. A Canadian study reviewed the quality of drug therapy guidelines developed or endorsed by Canadian organizations from 1994 to 1998.[14] Only 5% of the 217 guidelines reviewed met one half or more of the 20 criteria defining the rigor of development process. Another recent study assessed the quality of clinical practice guidelines in lung cancer with the AGREE instrument.[19] Of the 51 relevant guidelines identified, most were of poor overall quality. Only 19 of 51 (37%) of these guidelines were recommended for use in practice.

Application of pediatric guidelines

In this study, we recommended 14 of 17 (82%) pediatric, evidence-based, practice guidelines for use in the Netherlands. This high percentage of recommended guidelines is attributable in part to the fact that we searched deliberately for high-quality guidelines. We restricted our searches in MEDLINE and EMBASE with the term 'evidence-based', because searching MEDLINE for practice guidelines (publication type) yielded many consensus statements, position papers, workgroup reports, clinical policies, and standards. Although such documents are potentially useful in practice, their emphasis is usually not on the key clinical management questions and the supporting scientific evidence to answer those questions but rather on opinion and expert advice. We recognize that, because this was a selective search for potentially high-quality guidelines, we might have missed some good-quality guidelines.

All 14 guidelines were recommended with provisos or alterations. Typically, guideline recommendations are based on evidence, which is considered to be global, and other considerations, which are local and may differ among cultures. Each country has its own cultural and legal standards, and values and organizational limitations may affect the local recommendations. It is therefore not surprising that, in light of the same scientific evidence, different guideline developers produce different recommendations for local practice. For the successful implementation of good research evidence into clinical practice, we think that local groups of practitioners should create their own recommendations involving both all relevant evidence and all local stakeholders.

Sources of pediatric guidelines

The NGC is a freely available database with >1000 high-quality, evidence-based, practice guidelines. Not only American but also foreign guidelines can be registered if they fulfill a set of criteria, e.g. a systematic review of the literature should be performed. The SIGN and AAP guidelines are all registered in the NGC. The appraised guidelines indexed in the NGC and produced by the AAP were all recommended. Two of the 3 guidelines indexed in the CMA Clinical Practice Guidelines Infobase were not recommended because they scored low on rigor of development (2 and 11%). The guideline found in EMBASE was also not recommended; the score on rigor of development was 40% for this guideline.

Agencies that produce good-quality guidelines submit them to the NGC, which applies stringent quality criteria before it indexes a given guideline. To search for high-quality guidelines, therefore, we recommend starting by searching the NGC. However, in searching for a guideline on a specific topic that cannot be found in the NGC, it can be rewarding to search the Internet, because some guideline developers prefer simply to post their guidelines on their Web sites, as opposed to publishing them in journals or having them indexed in the NGC.

Areas for pediatric guideline improvement

On the basis of the results of our survey, we identified various areas in which pediatric guidelines can be improved. First, patient preferences and experiences were not sought. This is especially important for guidelines in which quality of life plays an important role or in which treatment can have significant morbidity or side effects. However, in the present selection of guidelines this might have been of lesser importance. Furthermore, most guidelines did not provide evidence of pilot testing. This is an important issue to ensure that the guideline can be put into actual clinical use.

Because we were looking for evidence-based guidelines, the AGREE domain scores for rigor of development are notable because they directly relate to how evidence based a guideline is. The fairly low mean score on this domain (54%) is partly due to the fact that the rigor of development domain contains not only questions on how evidence based a guideline is but also questions about whether the guideline was reviewed externally by experts before publication and whether a procedure for updating the guideline was provided. Items 8, 9 and 10 (i.e. whether systematic methods were used to search for evidence and whether the criteria for selecting the evidence and the methods used to formulate the recommendations were clearly described) are the most important items relating to the issue how evidence based a guideline is. If we only take these questions into account, the mean score would be 70% for our 14 recommended guidelines. It is generally recommended that guidelines be updated at least every 3 years, because new evidence can change the recommendations.[20] However, only 3 guidelines (18%) described a procedure for updating the guideline and 5 guidelines were already outdated by this criterion (i.e. >3 years old). For most recommended guidelines, we advised additional literature searches to update the evidence.

Another area in which most guidelines performed poorly was in the domain of applicability. Well-developed guidelines should include at least some consideration of potential barriers to implementation and cost implications, and they should supply monitoring criteria to assess the guideline's impact on practice organization and patient outcome. Finally, editorial independence was stated rarely. Poor performance in this domain could represent true conflicts of interest between funding sources and guideline development panels or might reflect simply poor reporting on these topics.

Limitations of the AGREE rating instrument

In this survey, we used the AGREE instrument to assess the quality of pediatric guidelines. Although this instrument is fairly new, it is one of the few guideline assessment tools with demonstrated validity and reliability.[16] A guideline that addresses the issues covered by the AGREE instrument is more likely to be a rigorously developed guideline. However, the AGREE instrument showed its limitations in this particular survey. We found that some of the variability in ratings might be attributable to differences in interpretation of several items. For example, for item 22, which is in the domain of edi-

torial independence, the observed agreement was 53%. This apparently poor agreement probably arises from the fact that some reviewers considered that the criterion was not met unless the statement was explicitly made in the guideline, whereas others interpreted the criterion to be met if the funding agency was a national government. In contrast, for three items (items 7, 8, and 17), the observed agreement was >90%. This is probably attributable to the fact that these questions are more straightforward. Another potential limitation of the AGREE instrument concerns the validity of the responses to the question on the overall assessment of the guideline's quality. No clear rules have been established as to how to weight the different domains. However, when reviewing the assessments compared with the domain scores, the responses appear to be valid and to reflect the quality of the guideline. For each guideline that was recommended by all 3 raters in the present survey, the overall domain scores were ≥50% for at least 3 domains, with an average of 4 domains. For guidelines that were not recommended, only 1 domain scored >50%. Furthermore, the scores on the domain rigor of development for recommended guidelines were high. All scores were ≥50%, with an average of 66%. Conversely, for guidelines that we did not recommend for use in our settings, the average score for this domain was only 18%. Another limitation of the AGREE instrument is that it assesses only the reporting of the different items and not the content validity of the recommendations. To asses the content validity of the recommendations, the rater must have both pediatric subject matter knowledge and skills in evidence-based medicine. When items are not reported specifically, they receive a low score (i.e. 1, strongly disagree), although in fact the developers might have met the criterion. It is therefore important that future guidelines report all different items specifically.

Opportunity for international collaboration on pediatric guidelines

The development of some parts of a guideline, specifically the comprehensive literature reviews, is time-consuming work. For some topics, we found several evidence-based guidelines with a substantial overlap in evidence summaries. Collaboration between guideline developers throughout the world could be a way to avoid unnecessary duplication of effort. One current collaboration in this respect is the International Guidelines Network (www.g-i-n.net), which has grown to 46 member organizations from 24 countries. The International Guidelines Network seeks to improve the quality of health care by promoting systematic development of clinical practice guidelines and their application into practice, by supporting international collaboration in guideline development.

Currently, no pediatric organization is member of the International Guidelines Network. If we had a worldwide collaboration in guideline development, then the work would be made lighter and faster through sharing agendas for topics for guideline development and state-of-the art guideline development methods, literature

searches, and critical appraisal. When a guideline is out of date, additional literature searches should be performed. Within a framework of guideline development collaboration local guideline development groups could produce their own recommendations without duplicating the time-consuming literature search and critical appraisal process. When, despite a rigorous search and analysis of the scientific literature, clear evidence for key recommendations is missing, it should be stated clearly on what basis local consensus has been reached. Also, recommendations for future empirical research to fill evidence gaps can be given.

Conclusions

The current volume of pediatric guidelines categorized as evidence based in popular databases is large. Compared with other studies of the quality of guidelines assessed with the AGREE instrument in fields outside pediatrics, these guidelines score well. This holds true especially for guidelines published and endorsed by the AAP or registered in the NGC. The AGREE instrument is a useful tool to select high quality guidelines from the international literature as candidates for adaptation to culture-specific values and local or national pediatric practices. However, to asses the content validity of the guideline and to decide on local applicability, users must have both pediatric subject matter knowledge and skills in evidence-based medicine. It is desirable to come to international collaboration among pediatric guideline developers, to exchange methodology of guideline development and evidence synthesis.

Acknowledgements

We thank Maruschka Merkus, PhD, for advice with the statistical analyses.

This project was funded by the Dutch Pediatric Society.

Appendix 1

AGREE Instrument

Scope and Purpose

1 The overall objectives of the guideline are specifically described.
2 The clinical questions covered by the guideline are specifically described.
3 The patients to whom the guideline is meant to apply are specifically described.

Stakeholder Involvement

4 The guideline development group includes individuals from all the relevant professional groups.
5 The patients' views and preferences have been sought.
6 The target users of the guideline are clearly defined.
7 The guideline has been piloted among end-users.

Rigor of development

8 Systematic methods were used to search for evidence.
9 The criteria for selecting the evidence are clearly described.
10 The methods used for formulating the recommendations are clearly described
11 The health benefits, side effects, and risks have been considered in formulating the recommendations.
12 There is an explicit link between the recommendations and the supporting evidence.
13 The guideline was externally reviewed by experts prior to its publication.
14 A procedure for updating the guideline is provided.

Clarity and Presentation

15 The recommendations are specific and unambiguous.
16 The different options for management of the condition are clearly presented.
17 Key recommendations are easily identifiable.
18 The guideline is supported with tools for application.

Applicability

19 The potential organizational barriers in applying the recommendations have been discussed.
20 The potential cost implications of applying the recommendations have been considered.
21 The guideline presents key review criteria for monitoring and/or audit purposes.

Editorial Independence

22 The guideline is editorially independent from the funding body.

23 Conflicts of interest of the guideline development members have been recorded.

Appendix 2

Title	Source
PULMONOLOGY / ALLERGY / EAR, NOSE, AND THROAT	
British Thoracic Society. *British guideline on the management of asthma*. Scottish Intercollegiate Guidelines Network; 2003	NGC, SIGN
National Asthma Education and Prevention Program. Expert Panel Report: *guidelines for the diagnosis and management of asthma:* Update on selected topics. National Asthma Education and Prevention Program; 1997 (revised 2002)	NGC, AAP (endorsed)
National Heart, Lung, and Blood Institute, World Health Organization. *Global initiative for asthma: global strategy for asthma management and prevention.* National Heart, Lung, and Blood Institute. National Heart, Lung, and Blood Institute, World Health Organization; 1995 (revised 2002)	NGC
Institute for Clinical Systems Improvement. *Diagnosis and management of asthma.* Institute for Clinical Systems Improvement; 1998 (revised 2000)	NGC
American Academy of Allergy, Asthma and Immunology/American College of Allergy, Asthma and Immunology/Joint Council of Allergy, Asthma and Immunology. *Practice parameters for the diagnosis and treatment of asthma.* American Academy of Allergy, Asthma and Immunology/American College of Allergy, Asthma and Immunology/Joint Council of Allergy, Asthma and Immunology; 1995 (reviewed 1998)	NGC
Boulet LP, Becker A, Berube D, Beveridge R, Ernst P. *Canadian Asthma Consensus Report*, 1999. Canadian Asthma Consensus Group. CMAJ. 1999 Nov 30;161(11 Suppl):S1-61	MEDLINE, CMA
Von Mutius E. Presentation of new GINA guidelines for paediatrics: *The Global Initiative on Asthma*. Clin Exp Allergy. 2000;30:6-10	MEDLINE
College of Physicians and Surgeons of Manitoba. Guidelines & statements: Guideline 913: *Treatment of acute asthma in children.* College of Physicians and Surgeons of Manitoba; 2000	CMA
Alberta Clinical Practice Guidelines Program. *Guideline for the management of acute asthma in adults and children.* Alberta Clinical Practice Guidelines Program; 1999	CMA
Children's Hospital Medical Center. *Evidence based clinical practice guideline for managing an acute exacerbation of asthma*. Cincinnati, OH: Children's Hospital Medical Center; 1999 (revised 2002)	NGC
Motala C, Kling S, Gie R, et al. *Guideline for the management of chronic asthma in children: 2000 update*. Allergy Society of South Africa Working Group. S Afr Med J. 2000;90:524-528, 530	MEDLINE, EMBASE
National Medical Research Council. *Management of asthma*. National Medical Research Council, Singapore Ministry of Health; 2002	NGC
Centers for Disease Control and Prevention. *Key clinical activities for quality asthma care*: recommendations of the National Asthma Education and Prevention Program. Centers for Disease Control and Prevention, National Asthma Education and Prevention Program; 2003	NGC
van Asperen Peter P, Mellis CM, Sly PD, Thoracic Society of Australia and New Zealand. *The role of corticosteroids in the management of childhood asthma.* Med J Aust. 2002;176:168-173	MEDLINE
American Association for Respiratory Care. *Selection of an aerosol delivery device for neonatal and pediatric patients.* American Association for Respiratory Care; 1995 (reviewed 2000)	NGC
American Association for Respiratory Care. *Selection of a device for the delivery of aerosol to the lung parenchyma.* American Association for Respiratory Care; 1996 (reviewed 2000)	NGC
Canadian Asthma Consensus Group. Canadian asthma consensus report, 1999: *Inhaled glucocorticosteroids in adults and children.* Canadian Asthma Consensus Group; 1999 (reviewed 2001)	CMA
American Association for Respiratory Care. *Infant/toddler pulmonary function tests.* American Association for Respiratory Care; 1995 (reviewed 2000)	NGC
American Association for Respiratory Care. *Static lung volumes: 2001 revision and update.* American Association for Respiratory Care; 2001	NGC

Title	Source
American Association for Respiratory Care. *Methacholine challenge testing: 2001 revision and update.* American Association for Respiratory Care; 2001	NGC
Finnish Medical Society Duodecim. *Prolonged cough in children*. Finnish Medical Society Duodecim; 2000 (revised 2001)	NGC
Children's Hospital Medical Center. *Evidence-based guidelines for the medical management of infants one year of age or less with a first time episode of bronchiolitis*. Cincinnati, OH: Children's Hospital Medical Center; 1996 (updated 2001)	NGC
Mandell LA, Marrie TJ, Grossman RF, Chow AW, Hyland RH. *Canadian guidelines for the initial management of community-acquired pneumonia: an evidence-based update* by the Canadian Infectious Diseases Society and the Canadian Thoracic Society: The Canadian Community-Acquired Pneumonia Working Group. Clin Infect Dis. 2000;31:383-421	MEDLINE
Children's Hospital Medical Center. *Evidence-based clinical practice guideline of community-acquired pneumonia in children 60 days to 17 years of age*. Cincinnati, OH: Children's Hospital Medical Center; 2000	NGC
Institute for Clinical Systems Improvement. *Viral upper respiratory infection (VURI) in children and adults.* Institute for Clinical Systems Improvement; 1994 (revised 2002)	NGC
American Academy of Allergy, Asthma and Immunology/American College of Allergy, Asthma and Immunology/Joint Council of Allergy, Asthma and Immunology. *Diagnosis and management of rhinitis*. American Academy of Allergy, Asthma and Immunology/American College of Allergy, Asthma and Immunology/Joint Council of Allergy, Asthma and Immunology; 1998	NGC
Allergic Rhinitis and its Impact on Asthma Workshop Group. *Allergic rhinitis and its impact on asthma*. Allergic Rhinitis and its Impact on Asthma Workshop Group; 2001	NGC
University of Michigan Health System. *Allergic rhinitis*. University of Michigan Health System; 2002	NGC
Infectious Diseases Society of America. *Practice guidelines for the diagnosis and management of group A streptococcal pharyngitis.* Infectious Diseases Society of America; 1997 (revised 2002)	NGC
Sinus and Allergy Health Partnership. *Antimicrobial treatment guidelines for acute bacterial rhinosinusitis*. Otolaryngol Head Neck Surg. 2000;123:5-31	MEDLINE
American College of Radiology. *Appropriateness Criteria for sinusitis in the pediatric population*. American College of Radiology; 1998	NGC
American Academy of Allergy, Asthma and Immunology/American College of Allergy, Asthma and Immunology/Joint Council of Allergy, Asthma and Immunology. *Parameters for the diagnosis and management of sinusitis.* American Academy of Allergy, Asthma and Immunology/ American College of Allergy, Asthma and Immunology/Joint Council of Allergy, Asthma and Immunology; 1998	NGC
Chan KH, Winslow CP, Levin MJ, et al. *Clinical practice guidelines for the management of chronic sinusitis in children*. Otolaryngology Head Neck Surgery. 1999;120: 328-334	EMBASE
American Academy of Pediatrics. *Clinical practice guideline: management of sinusitis.* American Academy of Pediatrics; 2001	NGC,AAP, MEDLINE
Children's Hospital Medical Center. *Evidence based clinical practice guideline for children with acute bacterial sinusitis in children 1 to 18 years of age.* Cincinnati, OH: Children's Hospital Medical Center; 2001	NGC
Children's Hospital Medical Center. *Evidence based clinical practice guideline for medical management of otitis media in children 2 months to 6 years of age.* Cincinnati, OH: Children's Hospital Medical Center; 1999	NGC
Alberta Clinical Practice Guidelines Program. *Guideline for the diagnosis and treatment of acute otitis media in children.* Alberta Clinical Practice Guidelines Program; 2000	CMA
Institute for Clinical Systems Improvement. *Diagnosis and treatment of otitis media in children*. Institute for Clinical Systems Improvement; 1995 (revised 2002)	NGC
American Academy of Pediatrics. *Managing otitis media with effusion in young children*. American Academy of Pediatrics; 1994 (reaffirmed 1997)	AAP, NGC
University of Michigan Health System. *Otitis media.* University of Michigan Health System; 1997 (revised 2002).	NGC

Title	Source
Donaldson LJ. Hayes JH. Barton AG. Howel D. Hawthorne M. *Impact of clinical practice guide lines on clinicians' behaviour: tonsillectomy in children.* Journal of Otolaryngology. 1999;28:24-30	EMBASE
American Academy of Allergy, Asthma and Immunology/American College of Allergy, Asthma and Immunology/Joint Council of Allergy, Asthma and Immunology. *Practice parameters for the diagnosis and management of immunodeficiency*. American Academy of Allergy, Asthma and Immunology/American College of Allergy, Asthma and Immunology/Joint Council of Allergy, Asthma and Immunology; 1995 (reviewed 1998)	NGC
American Academy of Allergy, Asthma and Immunology/American College of Allergy, Asthma and Immunology/Joint Council of Allergy, Asthma and Immunology. *Disease management of atopic dermatitis: a practice parameter.* American Academy of Allergy, Asthma and Immunology/ American College of Allergy, Asthma and Immunology/Joint Council of Allergy, Asthma and Immunology; 1997	NGC
INFECTIOUS DISEASES	
Swedish Consensus Group. *Management of infections caused by respiratory syncytial virus.* Scand J Infect Dis. 2001;33(5):323-328	MEDLINE
Infectious Diseases Society of America. *Quality standard for the treatment of bacteremia.* Infectious Diseases Society of America; 1994 (reviewed 1998)	NGC
Children's Hospital Medical Center. *Evidence based clinical protocol guideline for fever of uncertain source in infants 60 days of age or less*. Cincinnati, OH: Children's Hospital Medical Center; 1998	NGC
Children's Hospital Medical Center. *Evidence based clinical practice guideline of fever of uncertain source. Outpatient evaluation and management for children 2 months to 36 months of age*. Cincinnati, OH: Children's Hospital Medical Center; 2000	NGC
Kourtis AP. Sullivan DT. Sathian U. *Practice guidelines for the management of febrile infants less than 90 days of age* at the ambulatory network of a large pediatric health care system in the United States: summary of new evidence. Clinical Pediatrics. 2004;43:11-16	EMBASE
Canadian Paediatric Society. *Management of the febrile one- to 36-month-old child with no focus of infection.* Canadian Paediatric Society; 1996 (reviewed 2000)	CMA
Slater M. Krug SE. *Evaluation of the infant with fever without source: An evidence based approach*. Emergency Medicine Clinics of North America. 1999;17:97-126	EMBASE
Baraff LJ, Bass JW, Fleisher GR, et al. *Practice guideline for the management of infants and children 0 to 36 months of age with fever without source.* Agency for Health Care Policy and Research. Ann Emerg Med. 1993;22:1198-210. Erratum in: Ann Emerg Med 1993;22:1490	MEDLINE
Saiman L, Siegel J; Cystic Fibrosis Foundation. *Infection control recommendations for patients with cystic fibrosis: microbiology, important pathogens, and infection control practices to prevent patient-to-patient transmission.* Infect Control Hosp Epidemiol. 2003; 24(suppl):S6-52	MEDLINE
Health Canada, Laboratory Centre for Disease Control, Division of STD Prevention and Control. Canadian STD guidelines*: Human immunodeficiency virus (HIV) infection in children.* Health Canada, Laboratory Centre for Disease Control, Division of STD Prevention and Control; 1998	CMA
Canadian Paediatric Society. *HIV antibody testing in children.* Canadian Paediatric Society; 1993 (reviewed 2000)	CMA
Health Resources and Services Administration/National Institutes of Health/National Pediatric and Family HIV Resource Center. *Guidelines for the use of antiretroviral agents in pediatric HIV infection.* Health Resources and Services Administration/National Institutes of Health/National Pediatric and Family HIV Resource Center; 1999 (updated 2001)	NGC
British Columbia Centre for Excellence in HIV/AIDS. *Therapeutic guidelines for the treatment of HIV/AIDS and related conditions*: Part X, Section 6: Care of children with perinatal HIV infection. British Columbia Centre for Excellence in HIV/AIDS; 1998 (reviewed 1999)	CMA
Rachlis AR, Zarowny DP. *Guidelines for antiretroviral therapy for HIV infection*. Canadian HIV Trials Network Antiretroviral Working Group. CMAJ. 1998;158:496-505	MEDLINE
Canadian Paediatric Society. *Childhood tuberculosis : current concepts in diagnosis.* Canadian Paediatric Society; 1994 (reviewed 2000)	CMA
Infectious Diseases Society of America. *Practice guidelines for the treatment of tuberculosis*. Infectious Diseases Society of America; 2000	NGC

Title	Source
Canadian Paediatric Society. *Short course therapy for tuberculosis in infants and children.* Canadian Paediatric Society; 1994 (reviewed 2000)	CMA
Bass JB Jr, Farer LS, Hopewell PC, et al. *Treatment of tuberculosis and tuberculosis infection in adults and children*. American Thoracic Society and The Centers for Disease Control and Prevention. Am J Respir Crit Care Med. 1994;49:1359-1374	MEDLINE
Wormser GP, Nadelman RB, Dattwyler RJ, et al. *Practice guidelines for the treatment of Lyme disease.* Clin Infect Dis. 2000;31:1-14	NGC, AAP (endorsed)
Infectious Diseases Society of America. *Practice guidelines for the treatment of candidiasis*. Infectious Diseases Society of America; 2000	NGC
Canadian Paediatric Society. *Prevention and therapy of bacterial infections of the child with asplenia or hyposplenia.* Canadian Paediatric Society; 1999 (reviewed 2000)	CMA
American College of Critical Care Medicine. *Guideline for the management of intravascular catheter-related infections.* American College of Critical Care Medicine - Professional Association Infectious Diseases Society of America - Medical Specialty Society Society for Healthcare Epidemiology of America - Professional Association Society of Critical Care Medicine; 2001	NGC
Abzug MJ. *Meeting the challenge of antibiotic resistance: An evidence-based approach to treatment*. Pediatric Annals.1999;28:460-467	EMBASE
Canadian Paediatric Society. *Prophylactic antibiotics in children.* Canadian Paediatric Society; 1999 (reviewed 2000)	CMA
Singapore Ministry of Health. *Use of antibiotics in paediatric care*. Singapore Ministry of Health; 2002	NGC
Wojnarowska F, Kirtschig G, Highet AS, Venning VA, Khumalo NP; British Association of Dermatologists. *Guidelines for the management of bullous pemphigoid*. Br J Dermatol. 2002;147:214-221	MEDLINE
Kocher MS, Mandiga R, Murphy JM, et al. *Clinical practice guideline for treatment of septic arthritis in children: efficacy in improving process of care and effect on outcome of septic arthritis of the hip*. J Bone Joint Surg Am. 2003;85:994-999	EMBASE
NEPHROLOGY	
American Academy of Pediatrics. *Practice parameter: the diagnosis, treatment, and evaluation of the initial urinary tract infection in febrile infants and young children*. American Academy of Pediatrics, Committee on Quality Improvement, Subcommittee on Urinary Tract Infection. Pediatrics. 1999;103:843-852	MEDLINE, AAP, NGC
Children's Hospital Medical Center. *Evidence-based clinical practice guideline for patients 6 years of age or less with a first time acute urinary tract infection (UTI).* Cincinnati, OH: Children's Hospital Medical Center; 1999	NGC
Elder JS, Peters CA, Arant BS Jr, et al. *Pediatric Vesicoureteral Reflux Guidelines Panel summary report on the management of primary vesicoureteral reflux in children.* J Urol. 1997;157:1846-1851	MEDLINE
American Urological Association. *The management of primary vesicoureteral reflux in children.* American Urological Association; 1997	NGC, AAP (endorsed)
National Committee on Renal Care. *Glomerulonephritis.* National Committee on Renal Care, National Medical Research Council, Singapore Ministry of Health; 2001	NGC
National Kidney Foundation. *K/DOQI clinical practice guidelines for chronic kidney disease: evaluation, classification, and stratification*. National Kidney Foundation; 2002	NGC, MEDLINE
National Kidney Foundation. *Clinical practice guidelines for nutrition in chronic renal failure.* National Kidney Foundation; 2000	NGC
National Kidney Foundation. *Clinical practice guidelines for the treatment of anemia of chronic renal failure*. National Kidney Foundation; 1997 (updated 2000)	NGC
National Kidney Foundation. *Clinical practice guidelines for peritoneal dialysis adequacy.* National Kidney Foundation; 1997 (updated 2000)	NGC
National Kidney Foundation. *Clinical practice guidelines for hemodialysis adequacy.* National Kidney Foundation; 1997 (updated 2000)	NGC
Eknoyan G, Levin NW, Eschbach JW, et al. Continuous quality improvement: DOQI becomes K/DOQI and is updated: *National Kidney Foundation's Dialysis Outcomes Quality Initiative.* Am J Kidney Dis. 2001;37:179-194	MEDLINE

Title	Source
Canadian Paediatric Society. *Toilet learning: anticipatory guidance with a child-oriented approach.* Canadian Paediatric Society; 2000	CMA
GASTROENTEROLOGY	
Children's Hospital Medical Center. *Evidence-based clinical practice guideline for children with hypertrophic pyloric stenosis*. Cincinnati, OH: Children's Hospital Medical Center; 2001	NGC
Armon K, Stephenson T, MacFaul R, Eccleston P, Werneke U. *An evidence and consensus based guideline for acute diarrhoea management.* Arch Dis Child. 2001;85:132-142	MEDLINE
Infectious Diseases Society of America. *Practice guidelines for the management of infectious diarrhea*. Infectious Diseases Society of America; 2001	NGC
American Academy of Pediatrics. *Practice parameter: the management of acute gastroenteritis in young children*. American Academy of Pediatrics; 1996	AAP, NGC
Children's Hospital Medical Center. *Evidence-based clinical practice guideline for children with acute gastroenteritis*. Cincinnati, OH: Children's Hospital Medical Center; 1999 (revised 2001)	NGC
Canadian Paediatric Society. *Oral rehydration therapy and early refeeding in the management of childhood gastroenteritis.* Canadian Paediatric Society; 1994 (reviewed 2000)	CMA
Felt B, Wise CG, Olson A, et al. *Guideline for the management of pediatric idiopathic constipation and soiling.* Multidisciplinary team from the University of Michigan Medical Center in Ann Arbor. Arch Pediatr Adolesc Med. 1999;153:380-385	MEDLINE, EMBASE, NGC
Baker SS, Liptak GS, Colletti RB, et al. *Constipation in infants and children: evaluation and treatment.* A medical position statement of the North American Society for Pediatric Gastroenterology and Nutrition. J Pediatr Gastroenterol Nutr. 1999;29:612-626	NGC AAP (endorsed)
Canadian Helicobacter Study Group. Consensus conference on the *Approach to Helicobacter pylori Infection in Children and Adolescents.* Canadian Helicobacter Study Group; 1999	CMA
Gold BD, Colletti RB, Abbott M, et al. *Helicobacter pylori infection in children: recommendations for diagnosis and treatment*. J Pediatr Gastroenterol Nutr. 2000;31:490-497	MEDLINE, AAP (endorsed)
American Gastroenterological Association. *Medical position statement: celiac sprue.* American Gastroenterological Association; 2000 (reviewed 2001)	NGC
American College of Radiology. *Appropriateness Criteria for imaging recommendations for patients with Crohn's disease.* American College of Radiology; 1998 (revised 2001)	NGC
American Gastroenterological Association. *Medical position statement: guidelines for the evaluation of food allergies*. American Gastroenterological Association; 2000 (reviewed 2001)	NGC
American Gastroenterological Association. *Medical position statement: parenteral nutrition*. American Gastroenterological Association; 2001	NGC
North American Society for Pediatric Gastroenterology, Hepatology, and Nutrition. *Guidelines for evaluation and treatment of gastroesophageal reflux in infants and children.* North American Society for Pediatric Gastroenterology, Hepatology, and Nutrition; 2001	NGC
American Society of Health-System Pharmacists. *Therapeutic guidelines on stress ulcer prophylaxis*. American Society of Health-System Pharmacists; 1999	NGC
American College of Gastroenterology. *Diagnosis and management of achalasia*. American College of Gastroenterology; 1999	NGC
American Society for Parenteral and Enteral Nutrition. *Nutrition support for infants and children with specific diseases and conditions.* American Society for Parenteral and Enteral Nutrition; 1993 (reaffirmed 1999)	NGC
American Society for Parenteral and Enteral Nutrition. *Routes to deliver nutrition support in pediatric patients.* American Society for Parenteral and Enteral Nutrition; 1993 (reaffirmed 1999)	NGC
Anthony H, Collins CE, Davidson G, et al. *Pancreatic enzyme replacement therapy in cystic fibrosis: Australian guidelines:* Pediatric Gastroenterological Society and the Dietitians Association of Australia. J Paediatr Child Health. 1999;35:125-129	MEDLINE

Title	Source
NEUROLOGY	
Schutzman SA, Barnes P, Duhaime AC, et al. *Evaluation and management of children younger than two years old with apparently minor head trauma: proposed guidelines.* Pediatrics. 2001; 107:983-993	MEDLINE
Scottish Intercollegiate Guidelines Network. *Early management of patients with a head injury: a national clinical guideline*. Scottish Intercollegiate Guidelines Network; 2000	SIGN, NGC
American Academy of Family Physicians/American Academy of Pediatrics. *Management of minor closed head injury in children*. American Academy of Family Physicians/American Academy of Pediatrics; 1999	AAP, NGC
Children's Hospital Medical Center. *Evidence-based clinical practice guideline for management of children with mild traumatic head injury*. Cincinnati, OH: Children's Hospital Medical Center; 2000	NGC
Canadian Paediatric Society. *Management of children with head trauma.* Canadian Paediatric Society; 1990 (reviewed 1997)	CMA
National Committee on Neuroscience. *Head injury in children*. Singapore: National Committee on Neuroscience; 2001	NGC
Society of Critical Care Medicine. *Guidelines for the acute medical management of severe traumatic brain injury in infants, children, and adolescents*. Crit Care Med. 2003;31(suppl):S407-491	MEDLINE, EMBASE
College of Physicians and Surgeons of Manitoba. Guidelines & statements: Guideline 912: *Raised intracranial pressure in children.* College of Physicians and Surgeons of Manitoba; 1999	CMA
American Academy of Pediatrics. *Practice parameter: the neurodiagnostic evaluation of the child with a first simple febrile seizure*. American Academy of Pediatrics; 1996	AAP, NGC
American Academy of Pediatrics. *Practice parameter: long-term treatment of the child with simple febrile seizures*. American Academy of Pediatrics. Pediatrics. 1999;103:1307-1309	MEDLINE, AAP, NGC
Hirtz D, Ashwal S, Berg A, et al. *Practice parameter: evaluating a first nonfebrile seizure in children*: report of the quality standards subcommittee of the American Academy of Neurology, The Child Neurology Society, and The American Epilepsy Society. Neurology. 2000;55:616-623	MEDLINE
Hirtz D, Berg A, Bettis D, et al; *Practice parameter: treatment of the child with a first unprovoked seizure: r*eport of the Quality Standards Subcommittee of the American Academy of Neurology and the Practice Committee of the Child Neurology Society. Neurology. 2003;60:166-175	MEDLINE
Children's Hospital Medical Center. *Evidence-based clinical practice guideline for medical management of first unprovoked seizure in children 2 to 18 years of age.* Cincinnati, OH: Children's Hospital Medical Center; 1999 (revised 2002)	NGC
Appleton R, Choonara I, Martland T, et al. *The treatment of convulsive status epilepticus in children*. The Status Epilepticus Working Party, Members of the Status Epilepticus Working Party. Arch Dis Child. 2000;83:415-419	MEDLINE
Canadian Paediatric Society. *Therapy of suspected bacterial meningitis in Canadian children six weeks of age and older.* Canadian Paediatric Society; 2001	CMA
College of Physicians and Surgeons of Manitoba. Guidelines & statements: Guideline 911: *Lumbar punctures in children with suspected meningitis.* College of Physicians and Surgeons of Manitoba; 1997	CMA
Canadian Paediatric Society. *Acute childhood encephalitis and meningoencephalitis*. Canadian Paediatric Society; 1998 (reviewed 2000)	CMA
American College of Radiology. *Appropriateness Criteria for atraumatic isolated headache: when to image*. American College of Radiology; 1996 (revised 1999)	NGC
Shevell M, Ashwal S, Donley D, et al. *Practice parameter: evaluation of the child with global developmental delay*: report of the Quality Standards Subcommittee of the American Academy of Neurology and The Practice Committee of the Child Neurology Society. Neurology. 2003;60: 367-380	MEDLINE
ENDOCRINOLOGY	
American Diabetes Association. *Care of children with diabetes in the school and day care setting*. American Diabetes Association; 1998 (revised 2000; republished 2003)	NGC
American Diabetes Association. *Hyperglycemic crises in patients with diabetes mellitus*. American Diabetes Association; 2000 (revised 2001; republished 2003 Jan)	NGC

Title	Source
British Columbia's Children's Hospital. *A provincial protocol for the treatment of diabetic ketoacidosis in children.* British Columbia's Children's Hospital; 1998	CMA
College of Physicians and Surgeons of Manitoba. Guidelines & statements: Guideline 902*: Initial fluid and electrolyte management of diabetic ketoacidosis (DKA) in childhood.* College of Physicians and Surgeons of Manitoba; 1997	CMA
American Healthways. *Inpatient management guidelines for people with diabetes*. American Healthways; 1999	NGC
American Diabetes Association Task Force for Writing Nutrition Principles and Recommendations for the Management of Diabetes and Related Complications. *American Diabetes Association position statement: evidence-based nutrition principles and recommendations for the treatment and prevention of diabetes and related complications*. J Am Diet Assoc. 2002;102:109-118	MEDLINE
Scottish Intercollegiate Guidelines Network. *Management of diabetes*: *A national clinical guideline.* Scottish Intercollegiate Guidelines Network; 2001	NGC, SIGN
American Association of Clinical Endocrinologists/American College of Endocrinology. *Clinical practice guidelines for the evaluation and treatment of hyperthyroidism and hypothyroidism*. American Association of Clinical Endocrinologists/American College of Endocrinology; 1996	NGC
National Academy of Clinical Biochemistry. *Laboratory support for the diagnosis and monitoring of thyroid disease*. National Academy of Clinical Biochemistry; 2002	NGC
American Association of Clinical Endocrinologists/American College of Endocrinology. *Clinical practice guidelines for the diagnosis and management of thyroid nodules*. American Association of Clinical Endocrinologists/American College of Endocrinology; 1996	NGC
American Association of Clinical Endocrinologists/American College of Endocrinology. *Position statement on the prevention, diagnosis and treatment of obesity*. American Association of Clinical Endocrinologists/American College of Endocrinology; 1997 (revised 1998)	NGC
American Academy of Pediatrics. *Health supervision for children with Turner syndrome*. American Academy of Pediatrics; 2003	NGC, AAP
American Academy of Pediatrics. *Health supervision for children with Down syndrome.* American Academy of Pediatrics; 2001	NGC, AAP
American Academy of Pediatrics. *Evaluation of the newborn with developmental anomalies of the external genitalia.* American Academy of Pediatrics; 2000	NGC, AAP
ONCOLOGY / HEMATOLOGY	
Schiffer CA, Anderson KC, Bennett CL, et al. *Platelet transfusion for patients with cancer:* clinical practice guidelines of the American Society of Clinical Oncology. Journal of Clinical Oncology. 2001;19:1519-1538	EMBASE
Center for Disease Control and Prevention. *Guidelines for preventing opportunistic infections among hematopoietic stem cell transplant recipients*. Center for Disease Control and Prevention/Infectious Diseases Society of America/American Society for Blood and Marrow Transplantation; 2000	NGC, AAP (endorsed)
Infectious Diseases Society of America. *Guidelines for the use of antimicrobial agents in neutropenic patients with cancer*. Infectious Diseases Society of America; 1997 (revised 2002)	NGC, MEDLINE
American Society of Clinical Oncology. *Recommendations for the use of hematopoietic colony-stimulating factors: evidence-based, clinical practice guidelines*. J Clin Oncol. 1994;12:2471-2508	MEDLINE, NGC
Ozer H. *American Society of Clinical Oncology guidelines for the use of hematopoietic colony-stimulating factors*. Curr Opin Hematol. 1996;3:3-10	MEDLINE
American Society of Health-System Pharmacists. *Therapeutic guidelines on the pharmacologic management of nausea and vomiting in adult and pediatric patients receiving chemotherapy or radiation therapy or undergoing surgery.* American Society of Health-System Pharmacists;1999	NGC
American Society of Clinical Oncology. *Recommendations for the use of antiemetics: evidence-based, clinical practice guidelines*. American Society of Clinical Oncology; 1999	NGC
American Society of Anesthesiologists. *Practice guidelines for cancer pain management.* American Society of Anesthesiologists; 1996	NGC
De Mattia D, Del Principe D, Del Vecchio GC, et al. *Acute childhood idiopathic thrombocytopenic purpura: consensus guidelines for diagnosis and treatment.* Haematologica. 2000;85:420-424	EMBASE

Title	Source
American Society of Hematology. *Idiopathic thrombocytopenic purpura: a practice guideline developed by explicit methods for the American Society of Hematology*. American Society of Hematology; 1996	NGC
College of Physicians and Surgeons of Manitoba. Guidelines & statements: Guideline 920: *Sickle cell disease : early diagnosis and treatment of children.* College of Physicians and Surgeons of Manitoba; 1997	CMA
American Pain Society. *Guideline for the management of acute and chronic pain in sickle cell disease*. American Pain Society; 1999	NGC
Thrombosis Interest Group of Canada. *Thrombolytic therapy in children.* Thrombosis Interest Group of Canada; 2000	CMA
American College of Chest Physicians. *Antithrombotic therapy in children*. American College of Chest Physicians; 2001	NGC
Thrombosis Interest Group of Canada. *Warfarin therapy in children.* Thrombosis Interest Group of Canada; 2000	CMA
Thrombosis Interest Group of Canada. *Heparin in children*. Thrombosis Interest Group of Canada; 2000	CMA
Canadian Medical Association. *Guidelines for red blood cell and plasma transfusion for adults and children.* Canadian Medical Association; 1997	CMA
CARDIOLOGY	
Canadian Cardiovascular Society. Consensus Conference on Atrial Fibrillation: Part 10: *Atrial fibrillation and flutter in children and in young adults with congenital heart disease*. Canadian Cardiovascular Society; 1996	CMA
American College of Cardiology/American Heart Association. *Guidelines for the evaluation of management of heart failure***:** report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. American College of Cardiology/American Heart Association; 1995 (reviewed 2000)	NGC
Atkins DL, Dorian P, Gonzalez ER, et al. *Treatment of tachyarrhythmias*. Ann Emerg Med. 2001;37(4 Suppl): S91-109	MEDLINE
Cheitlin MD, Armstrong WF, Aurigemma GP, et al. *Guideline update for the clinical application of echocardiography:* summary article: a report of the American College of Cardiology/ American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol. 2003;42:954-970	MEDLINE
Babbs CF, Berg RA, Kette F, et al. American Heart Association; International Liaison Committee on Resuscitation. *Use of pressors in the treatment of cardiac arrest.* Ann Emerg Med. 2001; 37(4 Suppl):S152-162	MEDLINE
INTENSIVE CARE	
Babbs CF, Berg RA, Kette F, et al. *Use of pressors in the treatment of cardiac arrest*. Ann Emerg Med. 2001;37(4 Suppl):S152-162	MEDLINE
Chehardy P, Doherty A, Dracup K, et al. *Cardiopulmonary resuscitation and emergency cardiovascular care*. Ann Emerg Med. 2001;37(4 Suppl):S49-59	MEDLINE
American Association for Respiratory Care. *Selection of an oxygen delivery device for neonatal and pediatric patients*: *2002 revision and update*. American Association for Respiratory Care; 1996 (revised 2002)	NGC
American Association for Respiratory Care. *Management of airway emergencies.* American Association for Respiratory Care; 1995 (reviewed 2000)	NGC
Zideman D, Zaritsky A, Carlo W, et al. *Airways in pediatric and newborn resuscitation*. Ann Emerg Med. 2001;37(4 Suppl):S126-136	MEDLINE
Atkins DL, Chameides L, Fallat ME, et al. *Resuscitation science of pediatrics*. Ann Emerg Med. 2001;37(4 Suppl):S41-48	MEDLINE
Phillips B, Zideman D, Garcia-Castrillo L, Felix M, Shwarz-Schwierin V, European Resuscitation Council. *Guidelines 2000 for Advanced Paediatric Life Support*: A statement from Paediatric Life Support Working Group and approved by the Executive Committee of the European Resuscitation Council. Resuscitation. 2001;48:231-234	MEDLINE

Title	Source
Phillips B, Zideman D, Garcia-Castrillo L, Felix M, Shwarz-Schwierin U. European Resuscitation Council. *Guidelines 2000 for Basic Paediatric Life Support*: A statement from the Paediatric Life Support Working Group and approved by the Executive Committee of the European Resuscitation Council. Resuscitation. 2001;48:223-229	MEDLINE
European Resuscitation Council. *Pediatric basic life support*. European Resuscitation Council. Resuscitation. 2000;46:301-341	MEDLINE
Bingham R. *Evidence based paediatric resuscitation*. Paediatric Anaesthesia. 2001;11:1-2	EMBASE
Berg RA, Cobb LA, Doherty A, et al. American Heart Association; International Liaison Committee on Resuscitation: *Chest compressions and basic life support-defibrillation.* Ann Emerg Med. 2001;37(4 Suppl):S26-35	MEDLINE
NEONATOLOGY	
National Association of Neonatal Nurses. *Apnea of prematurity*. National Association of Neonatal Nurses; 1999	NGC
American Academy of Pediatrics/Canadian Paediatric Society. *Postnatal corticosteroids to treat or prevent chronic lung disease in preterm infants.* American Academy of Pediatrics/Canadian Paediatric Society; 2002	NGC
Niermeyer S, Van Reempts P, Kattwinkel J, et al. American Heart Association, International Liaison Committee on Resuscitation. *Resuscitation of newborns*. Ann Emerg Med. 2001; 37(Suppl):S110-125	MEDLINE
Wu TJ, Carlo WA, Neonatal Resuscitation Program. *Neonatal resuscitation guidelines 2000: framework for practice*. J Matern Fetal Neonatal Med. 2002 ;11: 4-10	NGC
Price D, International Guidelines 2000 Conference on Cardiopulmonary Resuscitation and Emergency Cardiac Care. *New neonatal resuscitation program guidelines*. Can Fam Physician. 2001;47:1263-1264	MEDLINE
Phillips B, Zideman D, Wyllie J, Richmond S, van Reempts P, European Resuscitation Council. *Guidelines 2000 for Newly Born Life Support.* A statement from the Paediatric Life Support Working Group and approved by the Executive Committee of the European Resuscitation Council. Resuscitation. 2001;48:235-239	MEDLINE
American Association for Respiratory Care. *Application of continuous positive airway pressure to neonates via nasal prongs or nasopharyngeal tube*. American Association for Respiratory Care; 1994 (reviewed 2000)	NGC
American Association for Respiratory Care. *Neonatal time-triggered, pressure-limited, time-cycled mechanical ventilation.* American Association for Respiratory Care; 1994 (reviewed 2000)	NGC
American Academy of Pediatrics. *Use of inhaled nitric oxide*. American Academy of Pediatrics; 2000	NGC, AAP
American Association for Respiratory Care. *Surfactant replacement therapy*. American Association for Respiratory Care; 1994 (reviewed 2000)	NGC
European Society of Cardiology. *Guidelines for the interpretation of the neonatal electrocardiogram.* European Society of Cardiology; 2002	NGC
American Academy of Pediatrics/Canadian Paediatric Society. *Prevention and management of pain and stress in the neonate*. American Academy of Pediatrics/Canadian Paediatric Society; 2000	CMA, NGC, AAP
Anand KJ. *Consensus statement for the prevention and management of pain in the newborn.* Arch Pediatr Adolesc Med. 2001;155:173-180	MEDLINE
Canadian Paediatric Society. *Guidelines for transfusion of erythrocytes to neonates and premature infants.* Canadian Paediatric Society; 1992 (reviewed 2000)	CMA
Nutrition Committee, Canadian Paediatric Society. *Nutrient needs and feeding of premature infants.* Nutrition Committee, Canadian Paediatric Society. CMAJ. 1995;152:1765-1785	MEDLINE
American Society for Parenteral and Enteral Nutrition. *Nutrition support for low-birth-weight infants*. American Society for Parenteral and Enteral Nutrition; 1993 Jul-Aug (reaffirmed 1999). 6 pages.	NGC
National Association of Neonatal Nurses. *Care of the neonate requiring parenteral nutrition*. National Association of Neonatal Nurses; 1999	NGC

Title	Source
College of Physicians and Surgeons of Manitoba. Guidelines & statements: Guideline 908: *Vitamin K deficiency in the neonate.* College of Physicians and Surgeons of Manitoba; 1999	CMA
American Academy of Pediatrics. *Practice parameter: management of hyperbilirubinemia in the healthy term newborn*. American Academy of Pediatrics; 1994 (reviewed 2000)	AAP, NGC
American Association for Respiratory Care. *Transcutaneous blood gas monitoring for neonatal and pediatric patients*. American Association for Respiratory Care; 1994 (reviewed 2000)	NGC
American Association for Respiratory Care. *Capillary blood gas sampling for neonatal and pediatric patients*. American Association for Respiratory Care; 1994 (reviewed 2000)	NGC
Evidence-based clinical practice guideline. *Neonatal skin care*. Association of Women's Health, Obstetric, and Neonatal Nurses - Professional Association. 2001 Jan. 54 pages.	NGC
OTHER	
Henderson-Smart DJ, Ponsonby A-L, Murphy E. *Reducing the risk of sudden infant death syndrome: A review of the scientific literature*. Journal of Paediatrics & Child Health. 1998;34:213-219	EMBASE
American Academy of Pediatrics. *Clinical practice guideline: diagnosis and management of childhood obstructive sleep apnea syndrome*. American Academy of Pediatrics; 2002	NGC, AAP, EMBASE
American Academy of Pediatrics. *Distinguishing sudden infant death syndrome from child abuse fatalities***;** American Academy of Pediatrics; 2001 (addendum published 2001)	NGC, AAP
American Academy of Pediatrics. *Apnea, sudden infant death syndrome, and home monitoring*. American Academy of Pediatrics; 2003	NGC, AAP MEDLINE,
Canadian Paediatric Society. *The child with multiple impairments.* Canadian Paediatric Society; 2000	CMA
American Academy of Pediatrics. *Clinical practice guideline: early detection of developmental dysplasia of the hip.* American Academy of Pediatrics; 2000	AAP, NGC
Canadian Paediatric Society. *Acetaminophen and ibuprofen in the management of fever and mild to moderate pain in children*. Canadian Paediatric Society; 1998 (reviewed 2000)	CMA
Murat I, Gall O, Tourniaire B. *Procedural Pain in Children: Evidence-Based Best Practice and Guidelines.* Regional Anesthesia & Pain Medicine. 2003;28:561-572	EMBASE
American College of Emergency Physicians. *Clinical policy for procedural sedation and analgesia in the emergency department.* American College of Emergency Physicians. Ann Emerg Med. 1998;31:663-677	MEDLINE
American Academy of Pediatrics. *Guidelines for monitoring and management of pediatric patients during and after sedation for diagnostic and therapeutic procedures*. American Academy of Pediatrics; 1992 (reaffirmed 1998; addendum published 2002)	NGC, AAP
American Academy of Pediatrics. *When inflicted skin injuries constitute child abuse*. American Academy of Pediatrics; 2002	NGC, AAP
American Academy of Ophthalmology. *Amblyopy.* American Academy of Ophthalmology; 1992 (revised 2002)	NGC
Singapore Ministry of Health. *Management of breastfeeding for healthy full-term infants.* Singapore Ministry of Health; 2002	NGC
American Academy of Pediatrics. *Identifying and treating eating disorders*. American Academy of Pediatrics; 2003	NGC, AAP
Thomas A, Forster G, Robinson A, Rogstad K, Clinical Effectiveness Group (Association of Genitourinary Medicine and the Medical Society for the Study of Venereal Diseases): *National guideline for the management of suspected sexually transmitted infections in children and young people*. Sex Transm Infect. 2002;78:324-331	MEDLINE

Reference List

1 Institute of Medicine. Guidelines for Clinical Practice: From Development to Use. Washington DC: National Academy Press; 1992

2 NICU study: best practices are not always followed. Healthcare Benchmarks Qual Improv 2003;10(6):68-69

3 Grol R. Improving the quality of medical care: building bridges among professional pride, payer profit, and patient satisfaction. JAMA 2001;286(20):2578-2585

4 Furdon SA, Eastman M, Benjamin K, Horgan MJ. Outcome measures after standardized pain management strategies in postoperative patients in the neonatal intensive care unit. J Perinat Neonatal Nurs 1998;12(1):58-69

5 Grimshaw JM, Russell IT. Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluations. Lancet 1993;342(8883):1317-1322

6 Kaempf JW, Campbell B, Sklar RS, Arduza C, Gallegos R, Zabari M et al. Implementing potentially better practices to improve neonatal outcomes after reducing postnatal dexamethasone use in infants born between 501 and 1250 grams. Pediatrics 2003;111 (4 Pt 2):e534-e541

7 Kuzma-O'Reilly B, Duenas ML, Greecher C, Kimberlin L, Mujsce D, Miller D et al. Evaluation, development, and implementation of potentially better practices in neonatal intensive care nutrition. Pediatrics 2003; 111(4 Pt 2):e461-e470

8 Perlstein PH, Kotagal UR, Bolling C, Steele R, Schoettker PJ, Atherton HD et al. Evaluation of an evidence-based guideline for bronchiolitis. Pediatrics 1999; 104(6):1334-1341

9 Turley K, Tyndall M, Roge C, Cooper M, Turley K, Applebaum M et al. Critical pathway methodology: effectiveness in congenital heart surgery. Ann Thorac Surg 1994; 58(1): 57-63

10 Wirtschafter DD, Jones KR, Thomas JT. Using health care outcomes to improve patient care in the NICU. Jt Comm J Qual Improv 1994; 20(2):57-65

11 Cluzeau FA, Littlejohns P, Grimshaw JM, Feder G, Moran SE. Development and application of a generic methodology to assess the quality of clinical guidelines. Int J Qual Health Care 1999; 11(1):21-28

12 Shaneyfelt TM, Mayo-Smith MF, Rothwangl J. Are guidelines following guidelines? The methodological quality of clinical practice guidelines in the peer-reviewed medical literature. JAMA 1999; 281(20):1900-1905

13 Ward JE, Grieco V. Why we need guidelines for guidelines: a study of the quality of clinical practice guidelines in Australia. Med J Aust 1996; 165(10):574-576

14 Graham ID, Beardall S, Carter AO, Glennie J, Hebert PC, Tetroe JM et al. What is the quality of drug therapy clinical practice guidelines in Canada? CMAJ 2001; 165(2):157-163

15 Grilli R, Magrini N, Penna A, Mura G, Liberati A. Practice guidelines developed by specialty societies: the need for a critical appraisal. Lancet 2000; 355(9198):103-106

16 AGREE collaboration. Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: the AGREE project. Qual Saf Health Care 2003; 12(1):18-23

17 Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. J Clin Epidemiol 1993; 46(5):423-429

18 Hoehler FK. Bias and prevalence effects on kappa viewed in terms of sensitivity and specificity. J Clin Epidemiol 2000; 53(5):499-503

19 Harpole LH, Kelley MJ, Schreiber G, Toloza EM, Kolimaga J, McCrory DC. Assessment of the scope and quality of clinical practice guidelines in lung cancer. Chest 2003; 123(1 Suppl):7S-20S

20 Shekelle PG, Ortiz E, Rhodes S, Morton SC, Eccles MP, Grimshaw JM et al. Validity of the Agency for Healthcare Research and Quality clinical practice guidelines: how quickly do guidelines become outdated? JAMA 2001; 286(12):1461-1467

21 Felt B, Wise CG, Olson A, Kochhar P, Marcus S, Coran A. Guideline for the management of pediatric idiopathic constipation and soiling. Multidisciplinary team from the University of Michigan Medical Center in Ann Arbor. Arch Pediatr Adolesc Med. 1999 Apr; 153(4):380-5

22 Baker SS, Liptak GS, Colletti RB, Croffie JM, Di Lorenzo C, Ector W, Nurko S. Constipation in infants and children: evaluation and treatment. A medical position statement of the North American Society for Pediatric Gastroenterology and Nutrition. J Pediatr Gastroenterol Nutr 1999 Nov;29(5):612-26

23 American Academy of Pediatrics. Committee on Quality Improvement. Subcommittee on Urinary Tract Infection. Practice parameter: the diagnosis, treatment, and evaluation of the initial urinary tract infection in febrile infants and young children. Pediatrics. 1999 Apr;103(4 Pt 1):843-52

24 Children's Hospital Medical Center (Cincinnati, OH). Evidence based clinical practice guideline for patients 6 years of age or less with a first time acute urinary tract infection (UTI). 1999. Available at: http://www.cincinnatichildrens.org/svc/dept-div/health-policy/ev-based/uti.htm. Accessed March 21, 2005

25 Schutzman SA, Barnes P, Duhaime AC, Greenes D, Homer C, Jaffe D, et al. Evaluation and management of children younger than two years old with apparently minor head trauma: proposed guidelines. Pediatrics. 2001 May;107(5):983-93

26 American Academy of Pediatrics, American Academy of Family Physicians. Management of minor closed head injury in children. Pediatrics 1999 Dec;104(6):1407-15

27 Scottish Intercollegiate Guidelines Network. Early management of patients with a head injury: A national clinical guideline. Edinburgh, Scotland: Scottish Intercollegiate Guidelines Network; 2000. SIGN publication 46. Available at: http://www.sign.ac.uk/guidelines/fulltext/46/index.html. Accessed March 21, 2005

28 Canadian Paediatric Society. Management of children with head trauma. CMAJ. 1990; 142(9): 949-952

29 Metzger DL. A provincial protocol for the treatment of diabetic ketoacidosis in children. B C Med J. 1998;40:110-114

30 Gralla RJ, Osaba D, Kris MG, et al. Recommendations for the use of antiemetics: evidence-based, clinical practice guidelines. American Society of Clinical Oncology. J Clinical Oncology. 1999;17:2971-2994

31 American Society of Health-System Pharmacists. ASHP Therapeutic guidelines on the pharmacologic management of nausea and vomiting in adult and pediatric patients receiving chemotherapy or radiation therapy or undergoing surgery. Bethesda, MD: American Society of Health-System Pharmacists;1999. Available at: http://www.ashp.org/bestpractices/index.cfm. Accessed March 21, 2005

32 American Urological Association. The management of primary vesicoureteral reflux in children. Linthicum, MD: American Urological Association; 1997

33 Baraff LJ, Bass JW, Fleisher GR, Klein JO, McCracken GH Jr, Powell KR, et al. Practice guideline for the management of infants and children 0 to 36 months of age with fever without source. Agency for Health Care Policy and Research [published correction appears in Ann Emerg Med. 1993 Sep;22(9):1490]. Ann Emerg Med. 1993 Jul;22(7): 1198-1210

34 Children's Hospital Medical Center. Evidence-based clinical practice guideline for fever of uncertain source in infants 60 days of age or less. Cincinnati, OH: Children's Hospital Medical Center; 1998. Available at: http://www.cincinnatichildrens.org/svc/dept-div/health-policy/ev-based/default.htm. Accessed March 21, 2005

35 Slater M. Krug SE. Evaluation of the infant with fever without source: An evidence based approach. Emergency Medicine Clinics of North America. Vol 17(1) (pp 97-126), 1999

36 Children's Hospital Medical Center. Evidence-based clinical practice guideline of fever of uncertain source: outpatient evaluation and management for children 2 months to 36 months of age. Cincinnati, OH: Children's Hospital Medical Center; 2000. Available at: http://www.cincinnatichildrens.org/svc/dept-div/health-policy/ev-based/default.htm. Accessed March 21, 2005

37 Canadian Paediatric Society. Management of the febrile one- to 36-month-old child with no focus of infection. Ottawa, Ontario, Canada: Canadian Paediatric Society; 1996 (reviewed 2000)

CHAPTER 8

General discussion and recommendations

During the last few years, our objective has been to practice Evidence-Based Pediatrics. Along the way, we encountered several major challenges, which will be discussed in this chapter. These challenges are 1) to use all existing evidence to guide clinical practice and the research agenda, 2) to avoid unnecessary duplication of effort, 3) to improve the quality of evidence in pediatrics, 4) to effectively translate research findings into pediatric clinical practice, and 5) to enhance the understanding of research methods among its end-users.

The discussion of these current challenges builds upon what we have learned in conducting the studies described in this thesis, integrated with knowledge that has been developed in this field during the last few years.

CHALLENGE 1: TO USE ALL EXISTING EVIDENCE TO GUIDE CLINICAL PRACTICE AND THE RESEARCH AGENDA

Rigorous child health systematic reviews

Several studies have shown that we do too much duplicative research that just confirms what is already established.[1] Frequently, more than one study has addressed the same clinical question. Because these studies enrolled different groups of people, in different settings, and used the interventions at hand differently in different trials, it is unlikely that they would all agree on the size and direction of the treatments' effects on patient health. The use of substandard study designs and methods may bias results as well. And indeed, studies on the same topic have often yielded quite opposite results, with the authors submitting conflicting conclusions. Thus, for both clinicians and patients it may be risky to base treatment decisions on a single study's results.

As an example in the area of prognostic research, the results of the individual studies included in our systematic review on neurodevelopmental outcome after neonatal hypoglycemia described in **Chapter 3** were quite conflicting: some studies found no differences between neonates with and without episodes of hypoglycemia on subsequent neurodevelopment, while others show serious brain damage after episodes of neonatal hypoglycemia. This explains the existing variation in clinical practice in the treatment of neonatal hypoglycemia. In **Chapter 3** we identified and discussed the several sources for this variability.

A rigorous and comprehensive systematic review is the most efficient way to access the information from all existing evidence on a given topic. Still, many reviews will have to be developed for common child health topics. The current challenge is: where to start? Clinicians and researchers interested in performing a systematic review could benefit from consulting resources such as the Cochrane Child Health Field (http://www.

cochranechildhealth.ualberta.ca). The mission of the Child Health Field is to ensure that children and adolescents receive effective interventions for the best health outcomes based on up-to-date evidence. The purpose of the Field is to support the conduct, dissemination and utilization of the child-relevant systematic reviews. The Child Health Field supports child-relevant Cochrane Review Groups through all steps in performing a systematic review. Users in the field have identified a list of *priority topics* in pediatrics, that need systematic reviews, and that The Field will encourage researchers and funders to consider (**Chapter 2**).

CHALLENGE 2: TO AVOID UNNECESSARY DUPLICATION OF EFFORT

Systematic reviews and evidence-based practice guidelines are seen as providing the best evidence to guide clinical practice and should be an integral part to the planning of future research.[2;3] They are time consuming to perform properly; valuable time busy clinicians lack. Therefore, before embarking on a new systematic review or an evidence-based practice guideline, researchers should be confident that the work has not already been done by others. Therefore it is crucial that child health systematic reviews and guidelines can be easily retrieved from existing bibliographic databases by both clinicians and researchers.

Child health systematic reviews

Where to find them?

In the Cochrane Library

The Cochrane Library is the main product of the Cochrane Collaboration, an international not-for-profit organization, providing up-to-date information about the effects of health care. The Cochrane Library consists of two databases that contain child health systematic reviews: The Cochrane Database of Systematic Reviews (CDSR) and The Database of Abstracts of Reviews of Effects (DARE). The CDSR contains Cochrane systematic reviews on health care interventions. The Cochrane Child Health Field keeps a list of all Cochrane reviews that are relevant to child health which can be found on their website: www.cochranechildhealth.ualberta.ca. The DARE database provides quality assessments of systematic reviews that have been published in journals identified by regular searches of important bibliographic databases such as MEDLINE, EMBASE, CINAHL, and by hand searching the major general journals. It does not include hand searched results from specific pediatric journals. Besides reviews on healthcare interventions, it contains systematic reviews on diagnostic test performance.

Several studies have shown that Cochrane reviews are more rigorous than journal reviews and they are updated frequently.[4;5] We recommend that busy clinicians, researchers and guideline developers searching for systematic reviews on child health care interventions should start searching The Cochrane Library (CDSR and DARE) for high quality systematic reviews. Finding systematic reviews in The Cochrane Library is fairly easy as they are already 'pre-filtered' in their focus on systematic reviews, so the only task is to find topic relevant systematic reviews in these databases.

In MEDLINE

In our study described in **Chapter 2** we have shown that DARE does only include about half of all child health systematic reviews published in peer reviewed pediatric journals. In 2004, DARE (The Cochrane Library, Issue 2, 2004: n=4645) included 431 child health systematic reviews of which 298 were also indexed in MEDLINE. If no relevant systematic reviews are found in The Cochrane Library or the interest is other than intervention or diagnosis, we recommend that busy clinicians search other databases such as MEDLINE next.

Finding systematic reviews in MEDLINE is challenging, as only a fraction of all citations in MEDLINE refer to (child health) systematic reviews and Medline's indexing procedures do not include 'systematic review' as a 'Publication Type'. To limit the search results from a query in MEDLINE, it is therefore necessary that a methodological filter be used. We assessed the usefulness of existing systematic review search filters, combined with a child filter, in finding child health systematic reviews in MEDLINE using the PubMed interface (**Chapter 2**).

Researchers conducting a new systematic review or guideline developers would best be served by the most sensitive search, Montori3. This search will have the highest probability of retrieving all existing relevant reviews, but will have low precision, retrieving many irrelevant articles. For a sensitive search Montori3 can be used in combination with the child filter (sensitivity: 94%, precision 3%). Those with little time on their hands, for example clinicians looking for answers to acute patient care questions, will likely be best served by a more precise search strategy – skipping many irrelevant articles. For a specific search Shojania's filter combined with the child filter can be used (sensitivity 74%, precision 45%). However, one in four child health systematic reviews will be missed by using this strategy.

Better reporting and definition of a systematic review

In our study described in **Chapter 2** it was often difficult to tell from the title and abstract if the article was a 'true' systematic review. In 1999 the Quality of Reporting Meta-analyses (QUOROM) statement was published to improve the quality of reporting meta-analyses or systematic reviews of clinical randomized controlled trials.[6] In order to be able to distinguish narrative reviews from systematic reviews, we recom-

mend that authors follow the recommendations given for reporting in the abstract. Otherwise, their reviews will not be picked-up by end-users in the field from the main bibliographical databases.

The number of records retrieved with sensitive systematic reviews filters was extremely high and in combination with a low precision, this means that a lot of irrelevant records are retrieved. These high number of records retrieved with the filters will put busy clinicians off searching for and eventually using the results from valuable systematic reviews. We recommend that minimal criteria that a review should fulfill in order for it to be indexed as a systematic review should be set up, making identification with PubMed easier.

The importance of a child health search filter

Ideally, a database of child health systematic reviews should be created. Pending this, we need validated 'child-health-filters' to be able to retrieve all child health relevant articles.

In **Chapter 2** we used a child filter, developed by the Cochrane Child Health Field, and added it to several existing systematic review filters to retrieve child health relevant systematic reviews more efficiently. Until recently, existing child filters have not been validated. We calculated the sensitivity of the child filter, because we wanted to be sure not to miss relevant child health systematic reviews. The sensitivity of the filter was excellent: 98% and precision increased with a factor 5-7. However, child filters need further validation, also in retrieving other child health relevant studies (e.g. RCTs).

Evidence-based pediatric guidelines

Where to find them?

There are several agencies known to produce and/or endorse evidence-based practice guidelines, e.g. the American Academy of Pediatrics (AAP: www.aap.org) and the Scottish Intercollegiate Guidelines (SIGN: www.sign.ac.uk). Relatively new are databases that register and link to evidence-based child health guidelines: the Agency for Health Care Policy and Research (AHCPR), the National Guideline Clearinghouse (NGC: www.guideline.gov) and the Canadian Medical Association Clinical Practice Guidelines Infobase (CMA: www.cma.ca).

In the National Guideline Clearinghouse

In our study described in **Chapter 7** we searched these databases to find pediatric evidence-based guidelines. In 2004, we found 215 pediatric evidence-based guidelines; most of the guidelines (n=119) were identified in the NGC. The National Guideline Clearinghouse (NGC) is a freely available database with more than 1,000 high quality evidence-based practice guidelines. Not only American, but also foreign guidelines can

be registered if they fulfill a set of criteria, e.g. a systematic review of the literature should be done. Guidelines produced by the AAP en SIGN are all indexed in the NGC. We found that the appraised guidelines indexed in the NGC all scored well on the AGREE instrument and were all recommended to use for pediatric clinical practice. In search for high quality guidelines, we therefore recommend to start searching for child health practice guidelines in the NGC.

Collaboration on guideline development

The development of some parts of an evidence based guideline, specifically the comprehensive literature review, is time consuming. Collaboration between guideline developers throughout the world could be a way to avoid unnecessary duplication of effort. One current collaboration in this respect is the 'International Guidelines Network', founded in 2002 and presently grown to 70 member organizations from 34 countries. In September 2006 more than 3,650 practice guidelines were available in their database (G-I-N: http://www.g-i-n.net).

G-I-N seeks to improve the quality of health care by promoting systematic development of clinical practice guidelines and their application into practice, through supporting international collaboration in guide development. Unfortunately, at present no pediatric organization is member of G-I-N. If national and international pediatric organizations joined such a worldwide collaboration in guideline development the work would be made lighter and faster through sharing agenda's for topics for guideline development and using state of the art guideline development methodology for literature searches and critical appraisal of available evidence. Within an international framework of guideline development collaboration local guideline developer groups could produce their own recommendations, without duplicating the time-consuming literature search- and critical appraisal-processes. Therefore, we recommend that pediatric organizations join G-I-N and foster international collaboration between *pediatric* guideline developers to exchange methodology of guideline development and evidence synthesis.

CHALLENGE 3: TO IMPROVE THE QUALITY OF EVIDENCE IN PEDIATRICS

The evidence-based medicine movement has brought about a paradigm shift not only in medical practice, but also in study design and in the appraisal of primary studies, systematic reviews and practice guidelines. Step 3 of practicing evidence-based medicine involves critical appraisal of the literature (see Introduction). For that purpose, scoring lists have been developed to assist the appraisal of primary studies (RCTs), diagnostic studies, prognostic studies and for systematic reviews and clinical practice guidelines (www.cochrane.nl).

Quality of child health systematic reviews

If clinicians are to have confidence that the results of systematic reviews can be used to guide clinical practice and the research agenda, then these reviews need to be of high quality. Unfortunately, most systematic reviews and meta-analyses published in peer reviewed journals have been shown to have methodological deficiencies that limit their validity and the applicability of their results in practice.[5;7-10] So far, these alarming results all come from studies on health care interventions in *adult* patients. Are results, in child health, any better?

In **Chapter 4** we present a study where we searched for all published systematic reviews on pediatric acute asthma management issues and evaluated the quality of these reviews. Methodological quality was assessed with the Overview Quality Assessment Questionnaire (OQAQ), currently the only validated instrument available for the critical appraisal of review articles.[11;12] We added items that are not covered by the OQAQ on the presence of clinical and statistical heterogeneity in reviews. A total of 23 systematic reviews were included, of which 14 had been published in the Cochrane Library and 9 in other peer reviewed journals. The overall quality according to the OQAQ was good, Cochrane reviews showing minimal flaws and journal reviews minor flaws (median scores 7 vs. 5; $P<0.001$). Although the reviews scored well on the OQAQ, we identified several areas limiting the usefulness of these systematic reviews for clinical practice. We will review these briefly, and reflect on avenues for improvement.

Source of funding

None of the systematic reviews stated the source of funding of the individual studies. Published trials sponsored by industry have been shown to exaggerate treatment effects and come to a positive conclusion in favor of the sponsor's drug five times more often than do not-for-profit-sponsored trials.[13] Like individual studies, reviews performed with industry sponsorship have been shown to have an increased risk of producing results in favor of the interventions promoted by the sponsors.[14;15] To enhance the correct interpretation of findings, we recommend that sources of funding of the included trials and of the review itself be clearly stated.

Separate reporting of children in mixed populations (adults and children) reviews

Another potential problem was that only 8 of the 15 reviews including both adults and children (so called mixed populations reviews) reported results for children separately. Mixed reviews present the problem of generalizability and applicability of study results: pediatricians may not be comfortable using mixed data results and applying them to the pediatric population. A recent study that investigated differences in effect sizes between adults and children in mixed populations systematic reviews could not ex-

clude clinically important differences due to lack of power.[16] Until there is clear evidence that there are no differences between adults and children in their response to the interventions evaluated we recommend that separate analyses for children be conducted.

Reporting of health outcomes

Many different health outcomes were reported by the systematic reviews included in our overview. Recently, new evidence has documented an additional threat to the validity of systematic reviews; selective reporting of trial outcomes within published studies.[17;18] The most common reason for the selective non-publication on health outcomes is lack of statistical significance. As selective reporting is widespread and can change the conclusions of systematic reviews we recommend that it should be routinely investigated in future systematic reviews.

Heterogeneity

Meta-analysis, i.e. the statistical pooling of results from individual studies, should only be considered when a group of trials is sufficiently homogeneous in terms of participants, interventions, health outcomes, and methods.[19] Statistical pooling of heterogeneous study results has been shown to produce serious errors in the estimation of effect sizes.[19;20]

First of all, the way reviews identified and assessed heterogeneity was quite diverse. Most of the individual studies included in the systematic reviews lacked clear definitions of included populations, making it hard to assess clinical heterogeneity. Despite this, the majority (87%) of the systematic reviews pooled data of individual studies. Most reviews used a statistical test to identify the presence of heterogeneity. However, these tests have low power to detect heterogeneity when there are small numbers of studies included.[20] Unfortunately, this is often the case in pediatric systematic reviews as the number of trials in children is limited. We believe that clinical and statistical heterogeneity should always be discussed and if present or difficult to assess due to a lack of description of included populations, we recommend that reviewers refrain from pooling. Furthermore, reviewers need guidance on how to cope with the various forms of heterogeneity in reviews.

Cochrane reviews versus journal reviews

Cochrane reviews scored better than journal reviews, especially in the review methods used and in reporting the results of the review. Journals could improve the quality and utility of systematic reviews they publish by providing authors and peer reviewers with clear reporting criteria and by encouraging authors to have the protocol checked by peer reviewers before the actual systematic review is performed – as the major general medical journals already do for clinical trials.[5]

Limitation of scoring lists for methodological quality

We found that the validity and usefulness of systematic reviews for clinical practice and in shaping the research agenda cannot solely be based on scoring lists like the OQAQ as important clinical details in individual trials may be overlooked. For instance, in our systematic review of systematic reviews on acute asthma management (**Chapter 4**), we found that most of the reviews scored well on methodological quality according to the OQAQ. However, we identified several areas limiting the validity of these reviews. One important problem was that the majority (83%) of the reviews did not state a clear definition of acute asthma, but instead used whatever definition was used by the included primary studies. Being pediatricians, we know that 'acute asthma' is not a clearly defined illness, but instead consists of patients with diverse clinical presentations. As treatments may have different effects sizes in patients with different clinical presentation of acute asthma (e.g. severe asthma), pooling of these potentially clinical heterogeneous studies results in an invalid overall effect size. The second problem is that without specifically stating population characteristics, clinicians will not be able to evaluate external validity, i.e. the extent to which the results are generalizable and applicable to their own population of children with acute asthma. Therefore, we believe that pediatricians with both clinical and methodological experience are needed on a team that authors systematic reviews.

Quality of evidence-based child health practice guidelines

The primary goal of child health practice guidelines is to improve the health of infants and children by ensuring they receive up to date evidence-based care. They are one of the various tools that could be used to improve the quality of care.[21] Several studies have shown that adherence to evidence-based guidelines leads to improvement in the quality of care provided.[22-27] But are child health guidelines of good enough quality and can we believe the recommendations presented?

Methodological quality: the AGREE Instrument

Over recent decades, the number of available clinical practice guidelines has grown enormously. Several studies highlighted concerns about the quality of clinical guidelines.[28-32] As the number of published guidelines proliferates there have been calls for the establishment of internationally recognized standards to improve the development and reporting of clinical guidelines. For this purpose an international group of researchers from thirteen countries, the Appraisal of Guidelines, Research and Evaluation (AGREE) Collaboration, has developed and validated a generic instrument that can be used to appraise the quality of clinical guidelines.[33]

In **Chapter 7** we set out to measure the volume of potentially high quality published pediatric guidelines and formerly assessed their quality with the AGREE instrument and see if we could adopt them for use in the Netherlands. A total of 215 guidelines were identified by the search and this list was sent to pediatricians in the Netherlands. They were asked to select the five most urgent topics for national pediatric guideline development. We appraised 17 existing evidence-based guidelines about the ten most frequently mentioned topics. Overall, these guides scored well compared to other studies on guideline quality in fields outside paediatrics, when assessed for quality with the AGREE instrument; after considering all domain scores, the reviewers recommended 14 of 17 (82%) guidelines to be used in local practice.

Content validity of child health practice guidelines recommendations

A limitation of the AGREE instrument is that it merely assesses the reporting of the different items, but not the content validity of the recommendations. For example, a guide can be developed according to AGREE standards but contain errors in the judgments of appraised articles, or, one can disagree with the 'other considerations' taken into account and influencing the final recommendations.

For example, in **Chapter 5** we describe the development of a pediatric evidence-based guideline on the first choice fluid for volume resuscitation. Existing evidence is poor (both in volume and quality) and therefore the final recommendations are mainly based on 'other considerations', in this case the potential side effects of colloids and crystalloids, current insight in pathophysiological mechanisms and their impact on the applicability of evidence from adults to children and neonates, and costs. And thus, to asses the content validity of the recommendations, the rater must have both pediatric subject matter knowledge and skills in evidence-based medicine. When items are not specifically reported, they will get a low score (namely '1; strongly disagree'), although in fact the developers may have met the criterion. It is therefore important that future guidelines report all different items specifically.

We advise future guideline developers to start searching for existing guidelines and appraise them with the AGREE instrument. If the guideline is to be considered 'recommended for use in local practice', additional literature searches should be performed to update on the evidence. For the successful implementation of a guideline into clinical practice we believe that local groups of practitioners and subspecialty content experts should create their own recommendations involving both all relevant evidence and all local stakeholders, as 'evidence is global, recommendations are local'.

Quality of primary studies

The Centre for Evidence-Based Medicine in Oxford regularly scans approximately 150 of the top medical journals in order to find good quality research articles to include in their bi-monthly journal (Evidence-Based Medicine). Only around 5% of the papers they find pass their basic criteria for methodological validity. However, of those 5%, only about 5% are actually clinically relevant. So, as a rule of thumb, only 5% of 5% (= 0.25%) of published research is worth basing practice on.[34]

For example, in our systematic review on the neurodevelopmental outcome after neonatal hypoglycemia (**Chapter 3**), we concluded that the methodological quality of the majority of the included 18 studies is poor and that none of the studies could validly quantify the effect of episodes of neonatal hypoglycemia on subsequent neurodevelopment. And so, recommendations for clinical practice could not be based on evidence because of a lack of valid empirical research.

Systematic reviews and clinical practice guidelines base their recommendations on the evidence generated by primary studies. Therefore, one of the most important challenges is to make sure that future studies in child health are of the highest methodological quality from the design stage to the reporting stage. We need to prevent that systematic reviews and clinical guidelines conclude that the results of a trial are probably biased or cannot be translated into practice for methodological reasons after it is completed and published. This effort will save money, time and, most importantly, our patient resources. In the following we discuss specific challenges that are related to primary study quality.

Addressing clinically relevant questions

Many trials fail to meet the needs of clinicians and patients because they are designed to meet the needs of pharmaceutical industry sponsors and specific national funding organizations. One reason why so many trials may not be particularly useful, is that they are designed to estimate the impact of an intervention under ideal circumstances in which the intervention is most likely to show benefit.[35;36] These *explanatory trials* relate to the efficacy of an intervention ('can it work'?). The outcomes that are assessed in the study are frequently not clinical outcomes but are often laboratory-based (surrogate) outcomes. *Pragmatic trials* are trials assessing the effectiveness ('does it work'?) of an intervention – the usefulness to real-world decision making. A pragmatic clinical trial would recruit patients from a variety of clinical settings with a range of clinical severities of a disorder and looking for clinical important outcomes. Most clinicians and patients will be interested in pragmatic trials showing the extent to which an intervention does more good than harm when provided under the usual circumstances. Therefore, both clinicians and patients should be involved in the design of such studies to define relevant treatment indications, patient populations, and the most important clinical outcomes.

Reducing the likelihood of bias

There are many sources of bias that can influence the results of RCTs. The two most important items that may lead to overestimation of treatment effect are a lack of allocation concealment and a lack of blinding of effect assessors.[37;38] However, many other sources of bias may occur in a clinical trial.

One of the problems with trying to assess the quality of a trial in terms of its methodological validity is that important elements of its design may not have been reported adequately, if at all. The 1996 guidelines called the CONSORT statement (Consolidation Standards of Reporting Trials) were presented for the reporting of RCTs.[39] For example, the CONSORT checklist stipulates that deviations in the protocol should be described including reasons for doing so. These standards have now been adopted by several leading journals and is now associated with improved reporting of RCTs.[40] However, even if a trial passes all criteria of CONSORT, the clinician should still consider the biological plausibility and applicability of the results to his own patients.

It has been empirically shown that reporting of trial outcomes is frequently incomplete, biased and inconsistent with protocols.[41;42] Overall, 50% of efficacy and 65% of harm outcomes per trial are incompletely reported. Statistically significant outcomes are reported more frequently compared with nonsignificant outcomes. This has potential major implications for clinical practice, as this selective outcome reporting bias will lead to overestimation of the effects of interventions in systematic reviews of the literature.

A well known problem is the influence of pharmaceutical industry sponsorship of trials.[43] Surveys have shown that manipulation of clinical trials – whereby, if the results are published at all, the control treatment is disadvantaged by design, analysis, or interpretation – is common.[41;43-45] Even when the results for the active and control therapies are no different, industry-sponsored trials come to a positive conclusion in favour of the sponsor's drug five times more often than do not-for-profit sponsored trials.[45] We recommend that governments and insurance companies save money and have patients treated in a more cost-effective manner by investing much more of their resources in (drug) trials than by relying on industry's own trials and conclusions. In this day and age it is a true challenge to get this message across to all buyers of care, policy makers, and funders of health care research.

In the same spirit, errors, research misconduct and bias in intervention research could be reduced markedly if ongoing initiatives to register all trials at their inception, ensuring public access to trial protocols and all data generated, become successful.[41;46] The International Committee of Medical Journal Editors have made a very positive and strong move towards this goal. They have agreed that after 1 July, 2005, its member journals, as a condition of considering a trial for publication, will require that it be registered in a public trials registry before patients enter the trial.[47] Publishing a RCT in one of the major general journals now requires that the protocol for the trial is sub-

mitted to the journal before the study starts and some journals even offer peer review of protocols.[48] One of the former editors of the British Medical Journal wrote an interesting essay in which he questioned if this approach would make a difference.[49] Studies funded by the industry are often found to be of good methodological quality according to scoring lists.[49;50] He believes that the companies seem to get the results they want not by 'fiddling with the results', but rather by asking the 'right' questions – and there are many ways to do so (Table 1). Therefore, and again, we need more public funding of trials, particularly of large head-to-head trials of all the treatments available for a specific condition.

Table 1 Examples of methods for pharmaceutical companies to get the results they want from clinical trials[49]

- Conduct a trial of your drug against a treatment known to be inferior
- Trial your drug against too low a dose of a competitor drug
- Conduct a trial of your drug against too high a dose of a competitor drug (making your drug seem less toxic)
- Conduct trials that are too small to show differences from competitor drugs
- Use multiple endpoints in the trial and select for publication those that give favorable results
- Do multicenter trials and select for publication results from centers that are favorable
- Conduct subgroup analyses and select for publication those that are favorable
- Present the results that are most likely to impress – for example, reduction in relative rather than absolute risk

Fostering collaboration in child health research

Given the smaller pool of patients available for trials in children, trials in children are usually inadequately powered to detect small to moderate treatment effects that might be of clinical significance.[51] To increase on sample size, we need large multi-centered, sometimes even multi-national trials. This will require efficient collaboration of pediatricians, both on a national and international level. We recommend that – to facilitate this process of collaboration in child health research – international scientific meetings be used to discuss study designs and protocols with all stakeholders in the field, and define patient populations, potential effect modifiers and clinically important outcomes. Another advantage of discussing study-protocols with all stakeholders in the field, is that eventual implementation of the results generated by that particular study will probably be easier – as they were involved in the design of the study.

Yet, we recognize that truly successful collaboration will not be easy, as it will involve important departures from current mechanisms of research funding and the rewarding of researchers. That it is feasible can be learned from 'mega-trials' being conducted in pediatric oncology research.[52-54]

Asking the right questions and answering them in a valid study

The true art is to ask the important research questions, do the systematic review to learn what is known on the topic and to identify knowledge gaps, design a valid study that will fill that gap and can actually be done in clinical reality, get the funding for it, and successfully work with collaborators to execute it. And to finally update the systematic review with the results of the new study.

By conducting the studies described in this thesis, we have practiced these steps to some extent. For example, based on what we learnt about the flawed studies included in our review on neonatal hypoglycemia and based on known pathophysiological concepts, we developed a set of truly relevant questions in this field. To answer the still open question about the long-term prognosis of neonatal hypoglycemia we proposed an optimal study design and we invited content experts and clinicians from all over the world to refine this design. This study was submitted to ZonMw (the Dutch MRC) and has been rewarded with a grant to perform a national multi-centre trial in the Netherlands starting in April 2007.

The role of conducting and compiling systematic literature reviews in this process cannot be stressed enough. We recommend that if the authors of a systematic review conclude that the methodological quality of existing studies is poor and more research is needed, the same authors provide the optimal design for such a future study building on the strengths and weaknesses of existing studies. Such a study design should then be discussed and refined with experts in the field.

CHALLENGE 4: TO EFFECTIVELY TRANSLATE RESEARCH FINDINGS INTO PEDIATRIC CLINICAL PRACTICE

On-going societal investment in basic and clinical research is ultimately only meaningful if this acquired knowledge is valid and translated into better care for patients. Unfortunately, in health care, there still is a substantial gap between what is known and what is actually done for patients.[55] It's not enough to rigorously show by a systematic review that an intervention works if we lack the power of implementing it, advertise it to the public, to physicians and to decision makers.

Many different approaches to knowledge translation have been advocated including development and dissemination of clinical practice guidelines. Clinical practice guidelines can facilitate translation of research into clinical practice and are seen as powerful tools to achieve effective care, reduce variability in daily practice, and may reduce costs.[56] Unfortunately, it is well known that many guidelines are not used in daily practice unless they are actively implemented pursuing consolidation of behavior change.[57;58] Quality-of-care indicators are used to determine failure or success of guideline implementation.[59;60]

In conducting the studies presented in this thesis we have learned several lessons related to the chance that research findings are actually translated into clinical practice through guidelines. Below we will focus on three key aspects in this respect, namely 1) conveying the strength of the research evidence underlying a guideline's recommendations to the end users, 2) the importance of a valid guideline development process, and 3) the art of effective practice guideline implementation.

The strength of research evidence underlying the recommendations

Users of evidence-based guidelines need to know how much confidence they can place in the recommendations, otherwise they will be reluctant to use the recommendations in their clinical practice. One of the challenges here is the following: In performing a systematic review or developing an evidence-based guideline, the reviewer must make numerous judgments about the available evidence. For many of these judgments, different reviewers could reasonably disagree about the status of the evidence. For example, one reviewer may view a certain methodological flaw as fatal, whereas another reviewer may view that same flaw as minor. This problem may be partially addressed through the use of standardized quality rating instruments. However, in the absence of empirical evidence on the extent to which a particular methodological flaw influences the results, the assessment of study quality necessarily entails judgment.

We acquired some experience with this. For example, to determine the methodological validity of the selected studies for our systematic review on neurodevelopmental outcome after neonatal hypoglycemia (**Chapter 3**), we used a widely quoted and internationally used checklist for the methodological quality of prognostic studies described by the Evidence-Based Medicine Working Group.[61;62] In order to overcome the problem of subjective judgment, we pre-specified each general methodological question from the list for our review and used a pilot sample to test the interrater agreement. We also specifically stated the methodological criteria we used to rate studies as 'high', 'moderate' or 'low' methodological quality.

Besides rating the methodological quality of evidence, other components of the strength of evidence are important, too: quantity of the evidence, consistency of results, and direction, magnitude and precision of determinant-outcome relationships in all evidence being considered.[63;64] For example, if there is only one small but well-conducted randomized controlled trial, how should one balance the high quality but low quantity of this evidence base to produce an overall strength-of-evidence rating? Or, how should one incorporate large treatment effects observed in studies with suboptimal designs into an overall strength-of-evidence rate?

This subjectivity of judgments at several points in the systematic review and guide-

line developmental process underscores the need for *transparency*. The end users should be able to decide for themselves whether the judgments used are reasonable. The GRADE working group has developed a rating system to assess the strength of a body of evidence.[65] The GRADE approach uses the primacy of study design (e.g. randomized trial, observational study) to set a starting quality grade. Then, other components, i.e. study quality (for example, allocation concealment, blinding), consistency (for example, similarity in results across studies) and directness (for example, generalizability of evidence), are considered which may increase or decrease the grade (Table 2).

Table 2 Overview of the GRADE system for grading the quality of evidence: criteria for assigning grade of evidence[65]

Criteria for assigning level of evidence	
Type of evidence	
Randomized trial	High
Observational study	Low
Any other type of research evidence	Very low
Increase level if	
Strong association	(+1)
Very strong association	(+2)
Evidence of a dose response gradient	(+1)
Plausible confounders reduced the observed effect	(+1)
Decrease level if	
Serious or very serious limitations to study quality	(-1) or (-2)
Important inconsistency	(-1)
Some or major uncertainty about directness	(-1) or (-2)
Imprecise or sparse data	(-1)
High probability of reporting bias	(-1)

The GRADE approach results in one of four outcome-specific grades: High, Moderate, Low, and Very Low (Table 3). The '*strength of a recommendation'* is then defined as the extent to which a clinician can be confident that adherence to the recommendation will result in greater benefit than harm for a patient. Their system for guiding complex judgments balances the need for simplicity with the need for full and transparent consideration of all important issues. Final recommendations involve a trade-off between benefits and harms, quality of evidence, applicability, and the uncertainty of the baseline risk.

Table 3 Overview of the GRADE system for grading the quality of evidence: definitions in grading the quality of evidence[65]

Level of evidence	**Definition**
High	Further research is not likely to change our confidence in the effect estimate
Moderate	Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate
Low	Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate
Very low	Any estimate of effect is uncertain

We recommend that future pediatric guideline developers use the GRADE system to grade the quality of evidence and the strength of recommendations.

Valid practice guideline development

There are well established key features of good clinical guideline programmes, which are summarized in Table 4.[66]

Table 4 Key features of good clinical guideline programmes[66]

People involved in guideline development
• credible organisation responsible for guideline development
• target users involved in guideline development ('ownership')
• balanced multidisciplinary guideline development group
• patient involvement at any stage of the development process
Methodology of guideline development
• systematic review of the literature, including existing high-quality guidelines
• combining evidence linkage and expert consensus in formulating recommendations
• external peer review
• formal update procedure
Dissemination and implementation strategies
• Production of different guideline formats, including patient versions, and tools for applications
• Use of the Internet
• Multiple implementation strategies

Once there is uncertainty about the most appropriate treatment a guideline should be developed according to evidence-based principles. The implementation should be considered part of the development process. Selection of topics, composition of the guideline development group, the work plan, search for evidence and involvement of clinical experts are all important in this.[67] We will briefly review these key ingredients here.

Topic selection

Appropriate topics can be selected by the relevance and prevalence of the problem, controversy about optimal care or existence of proven solutions.

Composition of the guideline group

Developing credible guidelines requires a balanced working group including clinical and methodological expertise to promote broad consensus and to prevent bias from conflicts of interest.

Formulating questions

Next, answerable questions around management are formulated on which the literature search will focus. Important are patient population, interventions and desired outcomes.

Reviewing the literature

The literature search starts by identifying existing guidelines and/or systematic reviews and appraise them for quality and applicability. If none are found, the next step if to perform a systematic review of the literature.

Formulation of recommendations

In formulating recommendations the scientific evidence and clinical expertise are brought together. The following issues should be considered: 1) Strength of the evidence (e.g. using the GRADE system); 2) Generalizability to the target population; 3) Cost-effectiveness of the proposed intervention; 4) Other considerations (e.g. patient preferences, adverse effects, pathofysiological mechanisms, etc).

Again, transparency is crucial, and therefore all 'other considerations' should be clearly presented in any guideline or its technical report. To reduce bias, a multidisciplinary group involving all stakeholders in the field should be involved in formulating the final recommendations. Finally an attractive, accessible format with brief summaries should be made available.[67;68]

We put these principles to the test. In **Chapter 5** we describe the methods and results of the guideline developmental process for the guideline 'Fluid resuscitation in neonatal and pediatric hypovolemic shock'. Several surveys have shown that, despite all evidence consistently showing that colloids are not superior to crystalloids[69-75], this evidence has not changed clinical practice: the majority of physicians still use costly colloid products.[76-78]

Before guideline development and implementation we found considerable variability in pediatric practice about the fluid of choice for volume resuscitation; about 50% of neonatologists and pediatric intensivists used colloids. In response to the uncertainty and variability in clinical practice regarding the optimal first choice fluid for volume resuscitation in children, we developed an evidence-based guideline, now endorsed by the Dutch Association of Pediatrics.

A systematic literature review was performed and all stakeholders were involved in formulating the guideline recommendations after two consensus meetings. They were asked to implement the guideline on their departments and their regional local hospitals. All members were visited by one pediatrician to identify barriers and facilitators for guideline implementation.

Barriers identified were addressed in the guideline e.g. the fear of inducing hypernatremia when using normal saline in neonates and applicability of evidence from

adults to children. Another barrier was the fact that international opinion leaders still advised to use albumin or synthetic colloids for hypovolemia. By clearly and explicitly showing what the existing evidence was to the stakeholders, we tried to convince them that these recommendations were not always based on evidence. The volume and quality of the available research evidence was limited. Therefore, the final recommendations are largely based on 'other considerations'. As these 'other considerations' may play a different role in other settings and countries, we explicitly described these 'other considerations'.

Pediatricians outside the Netherlands can decide for themselves if they agree with our process of decision making and make their own recommendations based on the information we have assembled.

Effective practice guideline implementation

Although we felt that the guide's recommendations were quite acceptable and could readily be translated into practice, we felt the need to add an active implementation phase to the process. In 2004 we received a grant from the Netherlands' Order of Medical Specialists to start a dedicated implementation project. In **Chapter 6** we present all implementation strategies that were used, and which indicators were used to evaluate the implementation success of the guideline on fluid resuscitation. We believe that from this project some important lessons can be learnt for future pediatric guideline development and implementation.

Involving stakeholders

One of the reasons guidelines are not used is that there has been mainly a focus on guideline development with little attention to implementation.[79] Implementation should start with development of the guideline and all stakeholders should be involved in the guideline development process. After guideline development and before active implementation, all (100%) neonatologists reportedly used normal saline as a first choice fluid for resuscitation and 88% of pediatric intensivists; albumin use for resuscitation had declined to zero. Our survey in a broader group of pediatric specialists showed that before active implementation most of the pediatric specialist's management behavior was already in keeping with the guideline's recommendations. We presume that stakeholders who were involved in the developmental process have been of great importance in disseminating the recommendations in their region.

Multifaceted interventions targeting identified barriers

Several systematic reviews including hundreds of studies have taught us that there is no single 'magic bullet' for implementation success.[55;57;58;80;81] This is not surprising as

the success of implementation is dependent on many factors and will therefore be guideline specific. Implementation strategies should be multifaceted: i.e. tailored to suit local circumstances and taking into account any particular barrier identified. Table 1 in **Chapter 6** presents the range of implementation strategies that we used. Strategies responded to the barriers to implementation that we had identified and included special incentives for guideline adopters.

Quality-of-care indicators

Success or failure of implementation should be measured by guideline-specific indicators.[59;60] We developed quality-of-care indicators consisting of reported behavior and data from medical records. Unfortunately, valid pharmaceutical data could not be obtained as fluid therapy in Dutch hospitals is not registered on indication or on number of patients and most often not separately for different departments. The health care sector should be made aware that in the process of improving quality of care appropriate and easily available measurements are urgently needed.

Study design for implementation studies

'Evidence-based implementation', which aims specifically to critically evaluate how summarized evidence can be best put into practice, is emerging as a critical area of research. There is a call for multiple arm cluster randomized controlled trials to investigate the (cost) effectiveness of different implementation strategies.[82;83] An important advantage of such designs is high internal validity. Disadvantages are the fact that they are expensive, time consuming, and generally will have low external validity and applicability in the case of comparing implementation strategies for guidelines. We decided not to perform a randomized trial comparing various different implementation strategies because we believe the results would not be applicable to other future guidelines' implementations, or be generalizable to other countries.

In a time where healthcare budgets are limited, we should carefully consider the most appropriate design of a study. First of all, the success of implementation is very much dependent on the guideline development process, i.e. the method of guideline development, strength of the underlying evidence, beliefs of the members of the guideline development group, etc. Secondly, the effect of different implementation strategies like educational meetings, outreach visits and the use of local opinion leaders, will all highly depend on the commitment and skills of the persons executing these strategies. Their effectiveness cannot be quantitated nor isolated as uncontrolled independent variables. As all these issues will be guideline specific, we decided it was more cost-effective and efficient to continuously develop and adjust implementation strategies according to identified barriers.

Learn from industry's marketing strategies

Several surveys have shown that fluid prescribing behavior is largely influenced by visits from drug detailers[76;77] and that physicians may change their prescribing practice as a result of contact with drug detailers, irrespective of scientific evidence.[84-86] Powerful industries have impressive control over dissemination and implementation of information. Marketing strategies by the industry involves regular visits of stakeholders and identifying barriers. We copied this approach in our implementation study and observed a marked change in fluid prescribing behavior even before active implementation. We believe we could learn from the industry's success of commercial marketing and must keep in mind that the industry spends huge amounts of money on marketing, on average five times more than they spent on developing new drugs. We believe that greater public investment is necessary in implementing unbiased research.

CHALENGE 5: TO ENHANCE THE UNDERSTANDING OF RESEARCH METHODS AMONG ITS END-USERS

In the end, even if primary studies, systematic reviews and clinical practice guidelines were perfectly designed, they could not influence health care decisions and outcomes if users could not understand their results. To correctly interpret an empirical study and subsequently translate the study's findings into pediatric clinical practice the end user needs to be able to critically appraise the methodological quality of the research. Most efforts to promote the understanding of research focus on researchers, particularly those in training. So far, little has been done to enhance the understanding of research methods among its end-users. If we want public money to be well spent, we have to develop and implement effective strategies to increase the understanding of research and the importance of its methodological quality to policymakers and all funders of medical research.

Conclusions

In conducting the studies described in this thesis, we have learned several lessons. We believe that the key elements in bridging the existing gap between research and practice in modern pediatrics are 1) **transparency** of the process of conducting primary research and preparing systematic reviews and practice guidelines, 2) **methodological quality** of the research, its syntheses and the practice guidelines based on these syntheses, and 3) **collaboration** between clinicians, guideline developers, researchers, funders, and patients. Table 5 summarizes all the lessons we have learnt in the form of recommendations for a) primary studies, b) systematic reviews, and c) clinical practice guidelines in 21st century pediatrics.

Table 5 Summary of Recommendations

A) PRIMARY STUDIES

Existing primary studies

- Should be summarized in methodologically rigorous systematic reviews (see below)
- Quality should be assessed using scoring lists for either randomized controlled trials, diagnostic or prognostic studies (www.cochrane.nl)
  - Be aware of 'good' reporting of 'bad' studies
  - Biological plausibility and applicability to your own patients should always be considered

Future primary studies

- Should be designed with highest methodological quality
- Should address clinically relevant questions
  - Clinicians and patients should be involved to define relevant patient populations and important clinical outcomes
- Should be reported according to CONSORT etc.
- Should be funded by independent sponsors
- Registration in prospective trial registers should be obligatory
- Collaboration should enhance the validity
  - Discuss study protocols with stakeholders in the field
  - Define patient populations, potential effect modifiers and clinically important outcomes
  - Multi-centre trials

B) SYSTEMATIC REVIEWS

Existing systematic reviews

- Where and how to find them?
  - Cochrane Library: CDSR and DARE
  - MEDLINE (PubMed interface) by using methodological filters
- Assessment of usefulness for clinical practice
  - By pediatricians with both clinical and methodological experience
    - * Use methodological scoring list: OQAQ
    - * Look for important clinical details, biological plausibility and applicability to your own patients
- A specialized child health systematic review database should be created
- Minimal criteria need to be set up that a review should fulfill in order for it to be indexed as a systematic review in large medical databases (e.g. MEDLINE)

Future systematic reviews

- Pediatricians with both clinical and methodological experience are needed to author systematic reviews
- Consult the Cochrane Child Health Field (www. cochranechildhealth.ualberta.ca) for support in conducting a systematic review
- Should be of excellent methodological quality
  - Use QUOROM for reporting of systematic reviews
  - Sources of funding of included trials and of the review itself should be clearly stated
  - Results on children should be separately analyzed and reported in mixed (adults and children) population reviews
  - Selective outcome reporting should be investigated
  - Reviewers need guidance on how to cope with heterogeneity

C) EVIDENCE-BASED CLINICAL PRACTICE GUIDELINES

Existing guidelines

- Where to find them?
  - National Guideline Clearinghouse
- Assessment of the quality of guidelines
  - By pediatricians with both clinical experience and methodological skills
  - Use a scoring list: the AGREE Instrument
  - Look for important clinical details, biological plausibility and applicability to your own patients

Future guidelines

- Future guideline developers should first search for existing guidelines and appraise them with the AGREE instrument. If recommended:
  - Additional updating literature searches should be performed
  - Local practitioners should formulate their own recommendations based on the evidence compiled in the existing guideline
- Collaboration between pediatric guideline developers is needed to exchange methodology of guideline development and evidence syntheses
- Guideline development
  - Start with implementation process right from the beginning
    - Involve all stakeholders (multidisciplinary group)
    - Identify implementation barriers and facilitators
    - Develop guideline specific quality-of-care indicators
  - Formulation of recommendations (be transparent)
    - Rating the strength of evidence (use GRADE system)
    - Other considerations should be clearly stated
    - Cost-effectiveness of the proposed interventions should be stated
- Guideline implementation
  - Use multifaceted interventions targeting identified barriers
  - Learn from the industry's success of marketing
  - More public investment is necessary

Reference List

1 Fergusson D, Glass KC, Hutton B, Shapiro S. Randomized controlled trials of aprotinin in cardiac surgery: could clinical equipoise have stopped the bleeding? Clin Trials 2005; 2(3):218-229.

2 Haines A, Jones R. Implementing findings of research. BMJ 1994; 308(6942):1488-1492.

3 Scholten RJ, Kremer LC. [Systemic reviews as a basis for guidelines]. Ned Tijdschr Geneeskd 2004; 148(6):262-264.

4 Jadad AR, Cook DJ, Jones A, Klassen TP, Tugwell P, Moher M et al. Methodology and reports of systematic reviews and meta-analyses: a comparison of Cochrane reviews with articles published in paper-based journals. JAMA 1998; 280(3):278-280.

5 Jadad AR, Moher M, Browman GP, Booker L, Sigouin C, Fuentes M et al. Systematic reviews and meta-analyses on treatment of asthma: critical evaluation. BMJ 2000; 320 (7234):537-540.

6 Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. Lancet 1999; 354(9193):1896-1900.

7 Assendelft WJ, Koes BW, Knipschild PG, Bouter LM. The relationship between methodological quality and conclusions in reviews of spinal manipulation. JAMA 1995; 274 (24):1942-1948.

8 Choi PT, Halpern SH, Malik N, Jadad AR, Tramer MR, Walder B. Examining the evidence in anesthesia literature: a critical appraisal of systematic reviews. Anesth Analg 2001; 92(3):700-709.

9 Delaney A, Bagshaw SM, Ferland A, Manns B, Laupland KB, Doig CJ. A systematic evaluation of the quality of meta-analyses in the critical care literature. Crit Care 2005; 9(5):R575-R582.

10 Kelly KD, Travers A, Dorgan M, Slater L, Rowe BH. Evaluating the quality of systematic reviews in the emergency medicine literature. Ann Emerg Med 2001; 38(5):518-526.

11 Oxman AD, Guyatt GH. Validation of an index of the quality of review articles. J Clin Epidemiol 1991; 44(11):1271-1278.

12 Oxman AD, Guyatt GH, Singer J, Goldsmith CH, Hutchison BG, Milner RA et al. Agreement among reviewers of review articles. J Clin Epidemiol 1991; 44(1):91-98.

13 Als-Nielsen B, Chen W, Gluud C, Kjaergard LL. Association of funding and conclusions in randomized drug trials: a reflection of treatment effect or adverse events? JAMA 2003; 290(7):921-928.

14 Barnes DE, Bero LA. Why review articles on the health effects of passive smoking reach different conclusions. JAMA 1998;279(19):1566-1570.

15 Jorgensen AW, Hilden J, Gotzsche PC. Cochrane reviews compared with industry supported meta-analyses and other meta-analyses of the same drugs: systematic review.

BMJ 2006; 333(7572):782.

16 Cramer K, Wiebe N, Moyer V, Hartling L, Williams K, Swingler G et al. Children in reviews: methodological issues in child-relevant evidence syntheses. BMC Pediatr 2005; 5:38.:38.

17 Chan AW, Hrobjartsson A, Haahr MT, Gotzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. JAMA 2004; 291(20):2457-2465.

18 Williamson PR, Gamble C, Altman DG, Hutton JL. Outcome selection bias in meta-analysis. Stat Methods Med Res 2005; 14(5):515-524.

19 Higgins JP, Green S e. Cochrane Handbook for Systematic Reviews of Interventions 4.2.5 [updated May 2005]. http:/www.cochrane.org/resources/handbook/hbook.htm.

20 Higgins J, Thompson S, Deeks J, Altman D. Statistical heterogeneity in systematic reviews of clinical trials: a critical appraisal of guidelines and practice. J Health Serv Res Policy 2002; 7(1):51-61.

21 Grol R. Improving the quality of medical care: building bridges among professional pride, payer profit, and patient satisfaction. JAMA 2001; 286(20):2578-2585.

22 Furdon SA, Eastman M, Benjamin K, Horgan MJ. Outcome measures after standardized pain management strategies in postoperative patients in the neonatal intensive care unit. J Perinat Neonatal Nurs 1998; 12(1):58-69.

23 Grimshaw JM, Russell IT. Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluations. Lancet 1993; 342(8883):1317-1322.

24 Kaempf JW, Campbell B, Sklar RS, Arduza C, Gallegos R, Zabari M et al. Implementing potentially better practices to improve neonatal outcomes after reducing postnatal dexamethasone use in infants born between 501 and 1250 grams. Paediatrics 2003; 111(4 Pt 2):e534-e541.

25 Kuzma-O'Reilly B, Duenas ML, Greecher C, Kimberlin L, Mujsce D, Miller D et al. Evaluation, development, and implementation of potentially better practices in neonatal intensive care nutrition. Paediatrics 2003; 111(4 Pt 2):e461-e470.

26 Perlstein PH, Kotagal UR, Bolling C, Steele R, Schoettker PJ, Atherton HD et al. Evaluation of an evidence-based guideline for bronchiolitis. Paediatrics 1999; 104(6):1334-1341.

27 Turley K, Tyndall M, Roge C, Cooper M, Turley K, Applebaum M et al. Critical pathway methodology: effectiveness in congenital heart surgery. Ann Thorac Surg 1994; 58(1): 57-63.

28 Cluzeau FA, Littlejohns P, Grimshaw JM, Feder G, Moran SE. Development and application of a generic methodology to assess the quality of clinical guidelines. Int J Qual Health Care 1999; 11(1):21-28.

29 Graham ID, Beardall S, Carter AO, Glennie J, Hebert PC, Tetroe JM et al. What is the quality of drug therapy clinical practice guidelines in Canada? CMAJ 2001; 165(2):157-163.

30 Grilli R, Magrini N, Penna A, Mura G, Liberati A. Practice guidelines developed by specialty societies: the need for a critical appraisal. Lancet 2000; 355(9198):103-106.

31 Shaneyfelt TM, Mayo-Smith MF, Rothwangl J. Are guidelines following guidelines? The methodological quality of clinical practice guidelines in the peer-reviewed medical literature. JAMA 1999; 281(20):1900-1905.

32 Ward JE, Grieco V. Why we need guidelines for guidelines: a study of the quality of clinical practice guidelines in Australia. Med J Aust 1996; 165(10):574-576.

33 Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: the AGREE project. Qual Saf Health Care 2003; 12(1):18-23.

34 Glasziou P. The EBM journal selection process. Evidence-Based Medicine 2006; 11(4):101.

35 Smyth RL, O'Byrne PM. The promise of and problems with evidence-based medicine for paediatric asthma management. Allergy 2002; 57 Suppl 74:5-14.:5-14.

36 Zwarenstein M, Oxman A. Why are so few randomized trials useful, and what can we do about it? J Clin Epidemiol 2006; 59(11):1125-1126.

37 Moher D, Pham B, Jones A, Cook DJ, Jadad AR, Moher M et al. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? Lancet 1998; 352(9128):609-613.

38 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(5):408-412.

39 Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I et al. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. JAMA 1996; 276 (8):637-639.

40 Plint AC, Moher D, Morrison A, Schulz K, Altman DG, Hill C et al. Does the CONSORT checklist improve the quality of reports of randomised controlled trials? A systematic review. Med J Aust 2006; 185(5):263-267.

41 Chan AW, Hrobjartsson A, Haahr MT, Gotzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. JAMA 2004; 291(20):2457-2465.

42 Schulz KF, Grimes DA. Multiplicity in randomised trials I: endpoints and treatments. Lancet 2005; 365(9470):1591-1595.

43 Gotzsche PC. Research integrity and pharmaceutical industry sponsorship. Trial registration, transparency and less reliance on industry trials are essential. Med J Aust 2005; 182(11):549-550.

44 Johansen HK, Gotzsche PC. Problems in the design and reporting of trials of antifungal agents encountered during meta-analysis. JAMA 1999; 282(18):1752-1759.

45 ls-Nielsen B, Chen W, Gluud C, Kjaergard LL. Association of funding and conclusions in randomized drug trials: a reflection of treatment effect or adverse events? JAMA 2003;290(7):921-928.

46 Chan AW, Upshur R, Singh JA, Ghersi D, Chapuis F, Altman DG. Waiving confidentiality for the greater good. BMJ 2006; 332(7549):1086-1089.

47 Clinical trial registration: a statement from the International Committee of Medical Journal Editors. Available at: http://www.icmje.org/clin_trial.pdf (accessed Dec 2006).

48 Jones G, Abbasi K. Trial protocols at the BMJ. BMJ 2004; 329(7479):1360.

49 Smith R. Medical journals are an extension of the marketing arm of pharmaceutical companies. PLoS Med 2005; 2(5):e138.

50 Lexchin J, Bero LA, Djulbegovic B, Clark O. Pharmaceutical industry sponsorship and research outcome and quality: systematic review. BMJ 2003; 326(7400):1167-1170.

51 Caldwell PH, Murphy SB, Butow PN, Craig JC. Clinical trials in children. Lancet 2004; 364(9436):803-811.

52 Chessells JM. Treatment of childhood acute lymphoblastic leukaemia: present issues and future prospects. Blood Rev 1992; 6(4):193-203.

53 Shuster JJ, Gieser PW. Meta-analysis and prospective meta-analysis in childhood leukemia clinical research. Ann Oncol 1996; 7(10):1009-1014.

54 Valsecchi MG, Masera G. A new challenge in clinical research in childhood ALL: the prospective meta-analysis strategy for intergroup collaboration. Ann Oncol 1996; 7(10):1005-1008.

55 Oxman AD, Thomson MA, Davis DA, Haynes RB. No magic bullets: a systematic review of 102 trials of interventions to improve professional practice. CMAJ 1995; 153(10): 1423-1431.

56 Grol R, Grimshaw J. From best evidence to best practice: effective implementation of change in patients' care. Lancet 2003; 362(9391):1225-1230.

57 Grimshaw JM, Shirran L, Thomas R, Mowatt G, Fraser C, Bero L et al. Changing provider behaviour: an overview of systematic reviews of interventions. Med Care 2001; 39(8 Suppl 2):II2-45.

58 Grimshaw JM, Thomas RE, MacLennan G, Fraser C, Ramsay CR, Vale L et al. Effectiveness and efficiency of guideline dissemination and implementation strategies. Health Technol Assess 2004; 8(6):iii-72.

59 Kuhlthau K, Ferris TG, Iezzoni LI. Risk adjustment for pediatric quality indicators. Paediatrics 2004; 113(1 Pt 2):210-216.

60 Lawrence M, Olesen F. Indicators of quality in health care. Eur J Gen Pract 1997; 3:103-108.

61 Altman DG. Systematic reviews of evaluations of prognostic variables. BMJ 2001; 323(7306):224-228.

62 Laupacis A, Wells G, Richardson WS, Tugwell P. Users' guides to the medical literature. V. How to use an article about prognosis. Evidence-Based Medicine Working Group. JAMA 1994; 272(3):234-237.

63 Bellomo R, Bagshaw SM. Evidence-based medicine: classifying the evidence from clinical trials – the need to consider other dimensions. Crit Care 2006; 10(5):232.

64 Treadwell JR, Tregear SJ, Reston JT, Turkelson CM. A system for rating the stability and strength of medical evidence. BMC Med Res Methodol 2006;6:52.:52.

65 Atkins D, Best D, Briss PA, Eccles M, Falck-Ytter Y, Flottorp S et al. Grading quality of evidence and strength of recommendations. BMJ 2004;328(7454):1490.

66 Burgers JS. Quality of clinical practice guidelines [PhD thesis]. Nijmegen University, 2002.

67 Wollersheim H, Burgers J, Grol R. Clinical guidelines to improve patient care. Neth J Med 2005; 63(6):188-192.

68 Grol R, Dalhuijsen J, Thomas S, Veld C, Rutten G, Mokkink H. Attributes of clinical guidelines that influence use of guidelines in general practice: observational study. BMJ 1998; 317(7162):858-861.

69 Human albumin administration in critically ill patients: systematic review of randomised controlled trials. Cochrane Injuries Group Albumin Reviewers. BMJ 1998; 317 (7153):235-240.

70 Alderson P, Schierhout G, Roberts I, Bunn F. Colloids versus crystalloids for fluid resuscitation in critically ill patients. Cochrane Database Syst Rev 2000;(2):CD000567.

71 Alderson P, Bunn F, Lefebvre C, Li WP, Li L, Roberts I et al. Human albumin solution for resuscitation and volume expansion in critically ill patients. Cochrane Database Syst Rev 2002;(1):CD001208.

72 Choi PT, Yip G, Quinonez LG, Cook DJ. Crystalloids vs. colloids in fluid resuscitation: a systematic review. Crit Care Med 1999; 27(1):200-210.

73 Finfer S, Bellomo R, Boyce N, French J, Myburgh J, Norton R. A comparison of albumin and saline for fluid resuscitation in the intensive care unit. N Engl J Med 2004; 350(22):2247-2256.

74 Schierhout G, Roberts I. Fluid resuscitation with colloid or crystalloid solutions in critically ill patients: a systematic review of randomised trials. BMJ 1998; 316(7136):961-964.

75 Velanovich V. Crystalloid versus colloid fluid resuscitation: a meta-analysis of mortality. Surgery 1989; 105(1):65-71.

76 Miletin MS, Stewart TE, Norton PG. Influences on physicians' choices of intravenous colloids. Intensive Care Med 2002; 28(7):917-924.

77 Schortgen F, Deye N, Brochard L. Preferred plasma volume expanders for critically ill patients: results of an international survey. Intensive Care Med 2004; 30(12):2222-2229.

78 Soderlind M, Salvignol G, Izard P, Lonnqvist PA. Use of albumin, blood transfusion and intraoperative glucose by APA and ADARPEF members: a postal survey. Paediatr Anaesth 2001; 11(6):685-689.

79 Massie J, Efron D, Cerritelli B, South M, Powell C, Haby MM et al. Implementation of evidence based guidelines for paediatric asthma management in a teaching hospital. Arch Dis Child 2004; 89(7):660-664.

80 Bero LA, Grilli R, Grimshaw JM, Harvey E, Oxman AD, Thomson MA. Closing the gap between research and practice: an overview of systematic reviews of interventions to promote the implementation of research findings. The Cochrane Effective Practice and Organization of Care Review Group. BMJ 1998; 317(7156):465-468.

81 Grimshaw J, Eccles M, Thomas R, MacLennan G, Ramsay C, Fraser C et al. Toward evidence-based quality improvement. Evidence (and its limitations) of the effectiveness of guideline dissemination and implementation strategies 1966-1998. J Gen Intern Med 2006; 21 Suppl 2:S14-20.:S14-S20.

82 Eccles M, Grimshaw J, Campbell M, Ramsay C. Research designs for studies evaluating the effectiveness of change and improvement strategies. Qual Saf Health Care 2003; 12(1):47-52.

83 Grimshaw J, Campbell M, Eccles M, Steen N. Experimental and quasi-experimental designs for evaluating guideline implementation strategies. Fam Pract 2000; 17 Suppl 1:S11-6.:S11-S16.

84 Chren MM, Landefeld CS. Physicians' behaviour and their interactions with drug companies. A controlled study of physicians who requested additions to a hospital drug formulary. JAMA 1994; 271(9):684-689.

85 Lurie N, Rich EC, Simpson DE, Meyer J, Schiedermayer DL, Goodman JL et al. Pharmaceutical representatives in academic medical centers: interaction with faculty and housestaff. J Gen Intern Med 1990; 5(3):240-243.

86 Orlowski JP, Wateska L. The effects of pharmaceutical firm enticements on physician prescribing patterns. There's no such thing as a free lunch. Chest 1992; 102(1):270-273.

Summary

This thesis is concerned with challenges that occur while practicing evidence-based pediatrics, i.e. the efficient and correct application of research evidence in modern pediatric clinical practice. The challenges addressed occur around the main EBM activities: searching for evidence, establishing the volume of the evidence, and determining the quality of the evidence. The thesis focuses on the three steps that are to be made between an original study as published in the medical scientific literature and clinical decision making in daily care. These steps are the summary of multiple studies into systematic reviews, the translation of these reviews into clinical practice guidelines, and the effective implementation of guidelines into current practice. All clinical topics presented in this thesis are 'question driven': they are topic areas of specific importance identified by practicing pediatricians.

Chapter 1 gives an introduction to the concepts of evidence-based pediatrics, systematic reviews and clinical practice guidelines.

Well conducted systematic reviews are currently seen as providing the best evidence to guide clinical practice. They are the basis for the recommendations of evidence-based practice guidelines, and should be an integral part of the planning of future clinical research. It is therefore that both clinicians and researchers should be able to reliably and quickly find existing systematic reviews. Bibliographic databases such as MEDLINE can be used to identify systematic reviews. Yet, finding systematic reviews in MEDLINE poses two challenges. First, only a fraction of all citations in MEDLINE are for systematic reviews. Second, MEDLINE's indexing procedures do not include 'systematic review' as a 'Publication Type'. To limit the search results from a query in MEDLINE, it is therefore recommended that a methodological filter be used, consisting of text words and MeSH headings directed to general indicators of systematic reviews in the MEDLINE record.

Chapter 2 describes a study in which the usefulness of nine existing systematic review search strategies to find child health systematic reviews in MEDLINE is evaluated by applying them, combined with a 'child filter', to a reference standard of child health systematic reviews. Data on sensitivity, number of records retrieved and precision are calculated in the domain of all current MEDLINE citations, and by combining filters with seven child health priority topics. The trade-off between sensitivity and precision will drive the choice of a filter. Researchers conducting a new systematic review and

guideline developers would best be served by the most sensitive search. This search will have the highest probability of retrieving all relevant reviews, but will have low precision, i.e. retrieving many irrelevant articles. Those with little time on their hands, e.g. clinicians looking for answers to patient care questions, will likely be best served by a more precise search strategy. We present the most sensitive filter (with relatively highest precision) and the most precise filter (with relatively highest sensitivity), combined with a child filter. Remarkably, the number of records retrieved with the different systematic review filters was in the (hundreds of) thousands and the precision was low. This will probably put busy clinicians off searching for and using the results generated by systematic reviews for their clinical decision making. One of the reasons for the low precision that we found is the fact that reporting of specific systematic review criteria in titles and abstracts was diverse. Reporting recommendations for the abstract as given by the QUOROM-statement should be used more strictly. In general medical databases, like MEDLINE, systematic reviews should ideally be indexed using a specific Publication Type term. However, since there is still no agreed definition of a systematic review, indexers will as yet not be able to introduce or apply such a term. We conclude that in order to make identification of systematic reviews using PubMed easier, there is an urgent need to set 'minimal' criteria that any review should fulfill in order for it to be indexed as a systematic review. Ideally, a database of child health systematic reviews should be created.

Hypoglycemia is the most common metabolic problem in neonatal medicine. Still, there is much controversy about the definition of a 'safe' blood glucose concentration, i.e. a value above which there will be no risk of long-term neurodevelopmental impairment. As a result of this controversy clinical practice varies widely. On a national basis, this leads to both over- and undertreatment of neonates. There is a great need to establish 'best practice'.

In **Chapter 3** we present a prognostic systematic review on neurodevelopment after episodes of neonatal hypoglycemia in the first week of life. Eighteen eligible studies were identified. The overall methodological quality of the included studies was considered poor in 16 studies and high in 2 studies. None of the studies provided a valid or precise estimation of the effect of neonatal hypoglycemia on neurodevelopment. We conclude that at present recommendations for clinical practice cannot be based on valid scientific evidence in this field. Building on the strengths and weaknesses of existing studies we proposed an 'optimal' future clinical study. This proposal was submitted to ZonMw (the Dutch MRC) and has been rewarded with a grant to perform a national multi-center trial comparing various management strategies for neonatal hypoglycemia in the Netherlands starting in April 2007.

An episode of acute asthma is a common reason for children to visit emergency departments. Various interventions exist for this prevalent condition. In this field, systematic reviews have already gained popularity as a way of coping with increasing amounts of information about new devices and drugs as they can synthesize large amounts of research evidence and help bridge the gap between evidence from single studies and clinical practice. Yet, if clinicians are to have confidence that the results of systematic reviews can be used to guide clinical practice and the research agenda, these reviews need to be of high quality.

Chapter 4 evaluates clinical, methodological and reporting aspects of systematic reviews on the treatment of acute asthma in children. Methodological quality was assessed with the Overview Quality Assessment Questionnaire (OQAQ) and with additional questions on clinical and statistical heterogeneity in these reviews. In addition, we assessed how data on children are reported in 'mixed population (i.e. adults and children)' reviews, and compared the quality of Cochrane reviews with non-Cochrane reviews published in peer reviewed journals. A total of 23 systematic reviews were included. The overall quality according to the OQAQ was good, Cochrane reviews showing *minimal flaws* and journal reviews *minor flaws* (median OQAQ scores 7 vs. 5; $P<0.001$). Yet, the usefulness for clinical practice is hampered by a lack of clear definitions of included populations, a lack of clinical important health outcomes, and a lack of separate reporting on children in mixed reviews. In addition we conclude that a major threat to these reviews' validity is the insufficient identification and handling of clinical and statistical heterogeneity. Recommendations for future trials and systematic reviews are given.

Considerable controversy and variation in clinical practice exists on whether colloids or crystalloids should be used for fluid resuscitation in pediatric hypovolemic shock. This uncertainty resulted in a plea from the profession for a Dutch evidence-based guideline.

Chapter 5 describes the development of the first national pediatric evidence-based guideline in the Netherlands on fluid resuscitation in pediatric hypovolemic shock. Twentynine stakeholders from across the Netherlands were involved in formulating the clinical questions for the guideline. A systematic literature review was performed and recommendations were formulated with the input of all stakeholders in two national consensus meetings. We found that in children the volume and quality of the available research evidence is limited. Therefore, the final recommendations are largely based on 'other considerations', in this case the potential side effects of colloids and crystalloids, current insight into pathophysiological mechanisms and their impact on the applicability of evidence from adults to children and neonates, and costs. The multidisciplinary guideline committee decided that isotonic saline should be the first choice fluid, because it is equally effective, safe and up to 100 times cheaper than albumin.

Practice guidelines are not used in daily practice unless actively implemented pursuing consolidation of health care workers' behavior change. Implementing guidelines is a great challenge and evidence about the most effective and efficient guideline implementation strategies in various different circumstances is lacking.

Chapter 6 describes the methodology used for the nationwide implementation of the evidence-based guideline on fluid resuscitation in pediatric hypovolemic shock described in Chapter 5, and reports the success of its implementation according to different indicators. We investigated fluid prescribing behavior for hypovolemia at three stages: 1) before guideline development in 2000; 2) after guideline development in 2004 and 3) after active implementation in a large sample of pediatric specialists in 2006, and identified potential barriers and facilitators for guideline implementation. Implementation success was evaluated using both questionnaires and data from medical records. In order to minimize costs and to optimize implementation effect, we continuously adjusted implementation strategies according to identified barriers. Before this guideline was developed there was controversy in the Netherlands about the fluid of choice for volume replacement: about 50% of neonatologists and pediatric intensivists used colloids. After guideline development but before the active implementation phase most of the pediatric specialist's management behavior was already in keeping with the guideline's recommendations: over 90% of the pediatricians used the recommended normal saline as a first choice fluid. Data from medical records were in keeping with reported behavior. We speculate that stakeholders who were involved in the developmental process have been of great importance in disseminating the recommendations in their hospital and in their region. To successfully implement clinical guidelines and reduce the enormous cost of active implementation, any guideline development team should consider the crucial aspects of implementation right from the beginning and involve all stakeholders in the developmental process. Implementation strategies should be targeting identified barriers. These strategies will therefore always be guideline specific. We question the ability of randomized controlled trials comparing various different implementation strategies to be extrapolated to other guideline implementation projects.

Anno 2007, pediatricians in the Netherlands are becoming more and more interested in using evidence-based practice guidelines. Yet, as the process of developing pediatric evidence-based guidelines is time-consuming and needs considerable resources and as the number of existing international pediatric evidence-based guidelines is rapidly increasing, the idea arose that we might adjust these guidelines for local use.

Chapter 7 describes the volume and quality of international evidence-based pediatric guidelines. Three pediatrician-reviewers appraised the available international guidelines on ten priority topics identified by Dutch pediatricians, using the Appraisal of Guidelines for Research and Evaluation (AGREE) instrument. A total of 215 evidence-

based pediatric guidelines were identified; of these, 17 guidelines on the 10 most urgent topics were appraised. Overall, these guidelines scored well when assessed for quality with the AGREE instrument; the reviewers recommended 14 of 17 guidelines (82%) to be used in local practice. This holds especially for guidelines published or endorsed by the American Academy of Pediatrics or registered in the National Guideline Clearinghouse. Furthermore, we give advise on how to adjust 'high quality' guidelines for local use. Typically, guideline recommendations are based on evidence, which is considered to be global, and other considerations, which are usually local and may differ among cultures. Each country has its own cultural and legal standards, and values and organizational limitations may affect the local recommendations. It is therefore no surprise that, in light of the same scientific evidence, different guideline developers produce different recommendations for local practice. We argue that for the successful implementation of good research evidence into clinical practice, local stakeholders should create their own recommendations involving both all relevant evidence and all local stakeholders. By using existing guidelines, the time-consuming literature search- and critical appraisal-processes do not have to be duplicated. To asses the content validity of the guideline and to decide on local applicability, users must have both pediatric subject matter knowledge and skills in evidence-based medicine. It is desirable to come to international collaboration among pediatric guideline developers, to exchange methodology of guideline development and evidence synthesis.

In **Chapter 8**, several current challenges in practicing Evidence-Based Pediatrics and Child Health are discussed. The discussion of these challenges builds upon what we have learned in conducting the studies described in this thesis, integrated with knowledge that has been developed in this field during the last few years. These challenges are:

1 To use all existing evidence to guide clinical practice and the research agenda
2 To avoid unnecessary duplication of effort
3 To improve the quality of evidence in pediatrics
4 To effectively translate research findings into pediatric clinical practice
5 To enhance the understanding of research methods among its end-users

From this discussion we conclude that the key elements in bridging the existing gap between research and practice in modern pediatrics are 1) **transparency** of the process of conducting primary research and preparing systematic reviews and practice guidelines, 2) **methodological quality** of the research, its syntheses and the practice guidelines based on these syntheses, and 3) **collaboration** between clinicians, guideline developers, researchers, funders, and patients. Finally, all the lessons we have learnt are translated into recommendations for future a) primary studies, b) systematic reviews, and c) clinical practice guidelines in 21st century pediatrics.

Samenvatting voor de leek

Introductie

Waarop baseren dokters hun beslissingen in de klinische praktijk? Het blijkt dat dit vaak niet berust op resultaten uit klinisch wetenschappelijk onderzoek. Er is dan ook veel variatie in de klinische praktijk van de kinderarts. Aan de ene kant worden bewezen effectieve behandelingen niet of onvoldoende toegepast en aan de andere kant worden mogelijk schadelijke behandelingen nog steeds gegeven.

Een eerste onderliggende oorzaak is dat artsen tegenwoordig overspoeld worden door medische informatie. De grootste en meest gebruikte elektronische database met klinische studies (MEDLINE) bevat inmiddels ruim 16 miljoen artikelen. Om op de hoogte te blijven van de ontwikkelingen voor de algemene kindergeneeskunde zou een kinderarts 5 artikelen per dag, 365 dagen per jaar moeten lezen.

Een tweede probleem is dat niet alle gepubliceerde onderzoeken van de benodigde hoge wetenschappelijke kwaliteit zijn en dat klinisch werkende artsen nog vaak onvoldoende toegerust zijn om dit te kunnen beoordelen. Bovendien zijn er meestal geen harde criteria waarop besloten kan worden of een onderzoek 'goed' of 'goed genoeg' is. Er zullen vele (deels subjectieve) afwegingen gemaakt moeten worden om tot een uiteindelijke beslissing voor de individuele patiënt te komen.

Om in deze toenemende informatiestroom van wisselende kwaliteit als arts niet ten onder te gaan is er een strategie bedacht: 'evidence-based medicine'. De definitie van evidence-based medicine is:

> 'het gewetensvol, expliciet en oordeelkundig gebruik maken van het huidige beste bewijsmateriaal om beslissingen te nemen voor individuele patiënten. De praktijk van evidence-based medicine impliceert het integreren van individuele klinische expertise met het beste externe bewijsmateriaal dat vanuit systematisch onderzoek beschikbaar is. De voorkeuren, wensen en verwachtingen van de patiënt spelen bij de besluitvorming een centrale rol'.

Het toepassen van evidence-based medicine in de klinische praktijk gaat volgens een vijfstapsmethode:

1 Het klinisch probleem vertalen in een beantwoordbare vraag
2 Het efficiënt zoeken naar het beste bewijsmateriaal
3 Het wegen van de gevonden evidence op methodologische kwaliteit en toepasbaarheid in de eigen praktijksituatie

4 Het nemen van beslissingen op grond van de beschikbare evidence
5 Het regelmatig evalueren van de kwaliteit van dit proces

Hoewel evidence-based medicine al in het begin van de jaren negentig vanuit Canada geïntroduceerd is, worstelen kinderartsen nu nog steeds met het toepassen van evidence-based medicine in de klinische praktijk. Een van de redenen is dat specifieke vaardigheden vereist zijn zoals het efficiënt kunnen zoeken naar medische literatuur en het kunnen beoordelen van de methodologische kwaliteit van allerlei verschillende studies. Over een bepaald onderwerp zijn vaak meerdere studies gepubliceerd, die eigenlijk alle gevonden en op methodologische kwaliteit beoordeeld moeten worden. Het kost dus tijd om evidence-based medicine te bedrijven; kostbare tijd waar drukke clinici doorgaans een gebrek aan hebben.

Voor het daadwerkelijk toepassen van evidence in de klinische praktijk zijn dan ook een aantal aanvullende stappen nodig. Allereerst is er behoefte aan uitgebreide samenvattingen van de literatuur betreffende een onderwerp. Voorheen gebeurde dit middels verhalende ('narrative') literatuuroverzichten, maar inmiddels hebben die plaats gemaakt voor *systematische* literatuuroverzichten (systematische reviews, SR). Een SR is gebaseerd op een expliciete vraagstelling, een zorgvuldige zoekstrategie naar de literatuur, een beoordeling van de kwaliteit van de onderzoeken, en een transparante presentatie van de resultaten. Naast een kwalitatieve samenvatting van de gegevens kan ook een kwantitatieve samenvatting (statistische pooling of meta-analyse) onderdeel zijn van een systematisch literatuuroverzicht. De voordelen zijn duidelijk: het proces is transparant, valide en reproduceerbaar – indien adequaat uitgevoerd. Een systematisch literatuuroverzicht is voor de behandelaar een efficiënte manier om snel een betrouwbaar inzicht te krijgen in alle bestaande evidence rond een klinische vraag.

Maar, dan zijn we er nog niet. De hoeveelheid evidence en de kwaliteit ervan zeggen ons namelijk nog niet wat we in de klinische praktijk moeten doen – of we een behandeling moeten toepassen. En soms is er geen (betrouwbaar) bewijs, maar zullen er in de praktijk toch beslissingen moeten worden genomen. Daarbij spelen 'andere overwegingen' een rol, zoals klinische relevantie (bijv. sterkte van het effect, consistentie van het bewijs, generaliseerbaarheid), veiligheid (bijv. bijwerkingen, risico's of complicaties op korte en lange termijn), patiëntenfactoren (bijv. specifieke behoeften van de patiënt, te verwachten therapietrouw), professioneel perspectief (bijv. tijdbesparing of verlies bij invoeren interventie, risico's voor de professional), beschikbaarheid van voorzieningen (bijv. in apparatuur, ervaring en deskundigheid), kosten(effectiviteit), etc. Deze 'andere overwegingen' worden meegenomen in Richtlijnen voor het Klinisch Handelen.

Deze Richtlijnen voor het Klinisch Handelen zijn bedoeld ter verbetering van de kwaliteit van zorg. Tot circa 1998 werden richtlijnen volgens de *consensusmethodiek*

opgesteld. Daarna is de methode *evidence-based* geworden, dat wil zeggen dat systematisch en uitgebreid aandacht wordt besteed aan het zoeken, selecteren, beoordelen en weergeven van literatuur ter onderbouwing van de aanbevelingen. Een *evidence-based richtlijn* is een document met aanbevelingen ter ondersteuning van de besluitvorming door professionals in de zorg en patiënten, en berusten op de resultaten van wetenschappelijk onderzoek met daarop gebaseerde discussie en aansluitende meningsvorming, gericht op het expliciteren van doeltreffend en doelmatig handelen.

Dit proefschrift

Dit proefschrift gaat over de uitdagingen die zich voordoen bij het toepassen van evidence-based medicine in de klinische kindergeneeskundige praktijk: het zoeken naar goede evidence, het vaststellen van de hoeveelheid evidence rond een bepaald klinisch probleem en de kwaliteit ervan. Het focus ligt op het bestaande 'gat' tussen het individuele klinische patiëntenonderzoek en de uiteindelijke beslissing in de klinische praktijk van de kinderarts. De belangrijkste elementen om dit gat te overbruggen zijn voor de kinderarts systematische literatuuroverzichten en evidence-based richtlijnen. Dit proefschrift gaat niet over een specifiek klinisch onderwerp, maar over verschillende, door kinderartsen geprioriteerde onderwerpen.

Zoals gezegd vormen systematische literatuuroverzichten de wetenschappelijke basis voor de aanbevelingen in richtlijnen. Zij behoren tot het hoogste niveau van bewijs voor het nemen van klinische beslissingen. Daarom is het van groot belang dat ze voor clinici en richtlijnmakers makkelijk te vinden zijn in elektronische medische databases. Er zijn gespecialiseerde databases voor systematische literatuuroverzichten, zoals de Cochrane Library. Echter, deze bevat niet alle bestaande literatuuroverzichten die voor de kinderarts relevant zijn. Een andere belangrijke bron voor systematische literatuuroverzichten is MEDLINE. Systematische literatuuroverzichten zijn hier moeilijker te vinden, omdat ze slechts een zeer klein percentage vormen van alle (16 miljoen) artikelen. Bovendien zijn de literatuuroverzichten niet als zodanig in MEDLINE geïndexeerd. De meest zinvolle manier om naar deze overzichten te zoeken is door zogenaamde zoekfilters te gebruiken, bestaande uit combinaties van methodologische termen die worden gebruikt in systematische literatuuroverzichten.

Hoofdstuk 2 beschrijft een onderzoek waar gekeken is naar de bruikbaarheid van bestaande SR zoekfilters voor het vinden van systematische literatuuroverzichten over kinderen in MEDLINE. Twee uitkomstmaten zijn hierbij van belang: 1) hoeveel van de in MEDLINE geïndexeerde systematische literatuuroverzichten worden daadwerkelijk met het filter gevonden (sensitiviteit) en 2) hoeveel van de gevonden artikelen betreffen daadwerkelijk een systematisch literatuuroverzicht (precisie). Om de precisie te ver-

hogen (en 'ruis' te verlagen) koppelden we aan het SR zoekfilter een kinderfilter – alleen systematische literatuuroverzichten die over ziekten bij *kinderen* gaan dienen te worden opgepikt. Richtlijnmakers zullen geïnteresseerd zijn in een sensitief filter: weinig van alle bestaande SRs zullen met deze zoekactie worden gemist, maar er worden wel meer irrelevante artikelen ('ruis') gevonden. Druk bezette kinderartsen zullen veelal geïnteresseerd zijn in een precies filter: weinig irrelevante artikelen worden opgepikt, maar iets meer SRs worden gemist.

Naar aanleiding van de resultaten van dit onderzoek kunnen aanbevelingen voor het beste sensitieve filter (met relatief hoogste precisie) en het beste precieze filter (met relatief hoogste sensitiviteit) worden geformuleerd. Deze filters kunnen gebruikt worden in combinatie met het kinderfilter om literatuuroverzichten bij kinderen te vinden in MEDLINE. Deze filters zijn echter verre van perfect. Het meest opvallende resultaat was dat het aantal gevonden artikelen ('hits') met de verschillende filters toch nog enorm hoog was: (honderd)duizenden hits, met daarbij een lage precisie, ofwel veel irrelevante artikelen. Wanneer hier een specifiek kindergeneeskundig klinisch onderwerp aan werd gekoppeld was het aantal hits beduidend minder. Toch moet een drukke kinderarts zich nog vaak door honderden tot duizenden referenties heen worstelen om uiteindelijk een paar relevante SRs te vinden. Dit zal de kinderarts ervan weerhouden deze overzichten te zoeken en dus te gebruiken in de klinische praktijk. De belangrijkste reden voor de vele (irrelevante) hits is dat de auteurs van SRs deze op zeer verschillende wijze rapporteren in het abstract en het zelfs vaak niet duidelijk wordt door het lezen van het abstract of het ook daadwerkelijk een SR betreft. Een ander probleem is dat er nog geen eenduidige definitie bestaat van een SR en ze dus ook niet als zodanig geïndexeerd kunnen worden in MEDLINE.

Tot slot komen we tot een aantal aanbevelingen. Idealiter zou er een aparte database moeten komen met daarin alle systematische literatuuroverzichten over kinderen. Het is van belang om internationaal tot consensus te komen over een aantal criteria waaraan een literatuuroverzicht minimaal moet voldoen om als een *systematisch* literatuuroverzicht geïndexeerd te worden. Ook is het van belang dat auteurs van SRs zich houden aan standaard afspraken voor het rapporteren van systematische literatuuroverzichten in het abstract zoals vastgelegd in QUOROM (Quality of Reporting of Meta-analysis).

Een lage bloedsuiker (hypoglycemie) is het meest voorkomende metabole probleem bij de pasgeborene; ongeveer 25% van alle pasgeborenen heeft een verhoogd risico op het ontstaan van een lage bloedsuiker. Omdat het kan leiden tot blijvende hersenbeschadiging en vertraagde psychomotore ontwikkeling, worden pasgeborenen met een hoog risico opgenomen, routinematig gescreend (geprikt) en zo nodig behandeld. Er is echter veel controverse over de hoogte van een 'veilige' glucose concentratie. Dit heeft tot gevolg dat de huidige klinische praktijk een grote variatie kent en leidt zowel tot

over- als tot onderbehandeling van pasgeborenen met hypoglycemie. Onder Nederlandse kinderartsen was er dan ook behoefte aan een evidence-based richtlijn. Als basis hiervoor hebben we een systematisch literatuuroverzicht gemaakt **(Hoofdstuk 3)** naar het effect van een lage bloedsuiker bij pasgeborenen in de eerste levensweek op de lange termijn psychomotore ontwikkeling.

We vonden 45 studies, waarvan 18 voldeden aan onze inclusiecriteria. Het merendeel van de studies (16/18) was van slechte methodologische kwaliteit, slecht 2 van de 18 van goede kwaliteit. Geen van de gevonden studies kon een uitspraak doen over het effect van episodes van hypoglycemie op de latere psychomotore ontwikkeling. Voor de klinische praktijk houdt dit in dat we op dit moment op basis van bestaande evidence geen aanbevelingen kunnen doen over bij welke bloedsuikerwaarde pasgeborenen met extra suiker behandeld dienen te worden om een ontwikkelingsachterstand op de lange termijn te voorkomen.

Er is dringend behoefte aan een goed opgezette en methodologisch valide studie die hier wel een uitspraak over kan doen. Lerend van de sterke en zwakke punten van de door ons beoordeelde studies, hebben we een optimale onderzoeksopzet ontworpen om deze studie te kunnen uitvoeren. Dit studievoorstel is inmiddels door ZonMw (Zorgonderzoek Nederland Medische wetenschappen) gehonoreerd met een subsidie voor 3 jaar. In april 2007 zal deze grote multicenter studie in Nederland van start gaan.

Acuut astma is een probleem waarmee veel kinderen zich melden op een eerste hulp van het ziekenhuis. Inmiddels zijn er vele systematische literatuuroverzichten verschenen naar de effecten van allerlei behandelingen bij kinderen met acuut astma. Deze literatuuroverzichten vormen voor de kinderarts een belangrijke samenvatting van alle evidence betreffende de effectiviteit van een bepaalde behandeling. Maar, het maken van deze systematische literatuuroverzichten moet aan bepaalde methodologische criteria voldoen willen ze valide zijn en dus bruikbaar voor de klinische praktijk.

Hoofdstuk 4 beschrijft een studie waar gezocht is naar alle bestaande systematische literatuuroverzichten over behandelingen bij kinderen met acuut astma. Deze overzichten zijn door twee (kinderarts-)onderzoekers met bestaande checklists (OQAQ: Overview Quality Assessment Questionnaire) beoordeeld op methodologische kwaliteit en op de bruikbaarheid voor de klinische praktijk.

In totaal vonden we 23 systematische literatuuroverzichten; 15 reviews includeerden studies bij zowel volwassenen als kinderen en 8 includeerden alleen studies bij kinderen. De methodologische kwaliteit van de systematische literatuuroverzichten was volgens de checklist over het algemeen goed. We identificeerden een aantal andere problemen die de bruikbaarheid van deze overzichten voor de klinische praktijk verminderen. Allereerst gaven de meeste SRs geen duidelijke definitie van de geïncludeerde populatie. Acuut astma is een ziektebeeld met een verscheidenheid aan klinische presentaties. Wanneer het niet duidelijk is welke patiënten onderzocht zijn, kunnen

we niet goed bepalen of de resultaten wel toepasbaar zijn op onze individuele patiënt op onze eerste hulp. Geen van de overzichten vermeldde of de geïncludeerde studies gesponsord werden door de industrie. Het is bekend dat door de industrie gesponsorde studies het effect van een behandeling overschatten. Een volgend probleem betrof het feit dat slechts 8 van de 15 literatuuroverzichten waar zowel volwassenen als kinderen werden geïncludeerd, kinderen ook separaat analyseerden en rapporteerden. Het is bekend dat behandelingen bij volwassenen een ander effect hebben dan bij kinderen. Bij astma zijn daar ook aanwijzingen voor. We kunnen daarom de resultaten van studies verricht bij volwassenen niet zo maar toepassen op kinderen. Daarom is de bruikbaarheid van dergelijke literatuuroverzichten voor de klinische praktijk gering. Een ander probleem is dat wanneer de afzonderlijke studies geïncludeerd in een review onderling te verschillend, te heterogeen zijn (bijv. ten aanzien van de onderzochte patiëntenpopulatie, interventie of uitkomstmaten), dan mogen de resultaten van deze studies niet gepoold worden om tot één effectschatter te komen. Zo´n effectschatter verkregen uit zeer divers onderzoek zou geen betekenisvolle waarde hebben en daarmee niet bruikbaar zijn voor de klinische praktijk. SRs moeten daarom het bestaan van deze heterogeniteit onderzoeken alvorens ze besluiten om de resultaten al dan niet te poolen. Wij vonden echter dat de systematische literatuuroverzichten zeer wisselend naar heterogeniteit keken en meestal toch overgingen tot poolen.

We adviseren dat toekomstige studies de geïncludeerde patiëntenpopulatie duidelijk beschrijven en dat er een internationaal geaccepteerde definitie komt van acuut astma. SRs die zowel studies bij volwassenen als kinderen includeren, moeten de studies bij kinderen apart analyseren en rapporteren. Reviewers hebben duidelijke voorschriften nodig hoe ze heterogeniteit tussen studies kunnen identificeren en hoe ze hier dan vervolgens mee om moeten gaan.

Een ander probleem waar de kinderarts in de klinische praktijk tegenaan loopt is welk vulmiddel het beste gegeven kan worden aan pasgeborenen en kinderen in zogenaamde circulatoire shock – een acuut tekort aan circulerend bloedvolume. Er zijn twee soorten vulmiddelen: colloïden en kristallijnen. Colloïden zouden effectiever zijn, omdat ze langer in de bloedbaan blijven; kristallijnen zijn het goedkoopst en hebben hoegenaamd geen bijwerkingen. In 1998 verschenen er enkele systematische literatuuroverzichten die lieten zien dat na het geven van een bepaald colloïd (albumine), er meer sterfgevallen waren dan na het geven van kristallijnen. Deze literatuuroverzichten hadden studies geïncludeerd bij volwassenen. Kinderartsen vroegen zich af of deze resultaten nu ook toepasbaar waren op de kindergeneeskundige populatie. Er was behoefte aan een evidence-based richtlijn.

In **Hoofdstuk 5** beschrijven we de door ons gebruikte methode voor het ontwikkelen van deze eerste Nederlandse evidence-based kindergeneeskundige richtlijn en de uiteindelijke aanbevelingen. We zijn begonnen met het formuleren van klinische vra-

gen. Vervolgens is er een systematisch literatuuroverzicht gemaakt over de mortaliteit en morbiditeit na het geven van colloïden versus kristallijnen bij pasgeborenen en kinderen in circulatoire shock. Er waren onvoldoende (goede) studies bij kinderen om hierover een uitspraak te kunnen doen. Maar, geen van de studies liet een voordeel zien van de duurdere colloïden ten opzichte van de goedkopere kristallijnen. Om tot de aanbevelingen voor de klinische praktijk te komen hebben we een multidisciplinaire klankbordgroep samengesteld bestaande uit 29 specialisten (kinderartsen, intensive care artsen, neonatologen, verpleegkundigen) om op basis van het wetenschappelijke bewijsmateriaal tot uiteindelijke aanbevelingen te komen voor de klinische praktijk. Alle 'andere overwegingen' die hierbij een rol hebben gespeeld zijn expliciet genoemd in de richtlijn. Dit waren pathofysiologische overwegingen, toepasbaarheid van resultaten uit volwassen studies bij kinderen, en kosten van de vulmiddelen. Voor de gebruikers van de richtlijn is het dus helder hoe de aanbevelingen tot stand zijn gekomen.

De richtlijn adviseert om bij pasgeborenen en kinderen met een shock fysiologisch zout te geven als eerste keus vulmiddel.

Uit gedegen onderzoek is het inmiddels bekend dat het verspreiden van richtlijnen onder artsen niet of onvoldoende leidt tot het veranderen van de klinische praktijk. Verschillende implementatie- (ofwel invoerings)strategieën zijn noodzakelijk om ervoor te zorgen dat een richtlijn ook daadwerkelijk gebruikt wordt. Effectieve implementatie van een richtlijn vereist een systematische aanpak en begint al bij het ontwikkelen van de richtlijn.

In **Hoofdstuk 6** wordt beschreven welke implementatiestrategieën gebruikt zijn om de richtlijn 'volumesuppletie' landelijk in de kindergeneeskundige praktijk te implementeren en wat het effect was op het voorschrijfgedrag van kinderartsen ten aanzien van de verschillende vulmiddelen. Onze eerste strategie was het betrekken van de doelgroep (via de multidisciplinaire klankbordgroep) bij het maken van de richtlijn. Alle leden van deze groep zijn persoonlijk bezocht door de projectleider om belemmerende en bevorderende factoren te identificeren welke een rol kunnen spelen bij het invoeren van de richtlijn in de praktijk. Op basis van deze belemmerende en bevorderende factoren zijn meerdere implementatiestrategieën ontwikkeld. Vervolgens is een systematisch literatuuroverzicht gemaakt van de bestaande evidence. Tijdens twee klankbordgroepvergaderingen zijn op basis van de evidence de aanbevelingen geformuleerd. Alle 'andere overwegingen' die hierbij een rol hebben gespeeld zijn expliciet genoemd in de richtlijn. Om de veranderingen in de praktijk ook te kunnen meten, zijn richtlijnspecifieke kwaliteitsindicatoren ontwikkeld. Het liefst hadden we apotheekgegevens geëvalueerd betreffende de hoeveelheid gebruikte colloïdale oplossingen en kristallijne oplossingen (in liters) vóór het maken van de richtlijn, na het maken van de richtlijn, en na de actieve implementatiefase. Helaas waren deze gegevens niet beschikbaar. Daarom is door middel van enquêtes nagegaan wat kinderartsen als hun

eerste keus vulmiddel aangaven vóór het maken van de richtlijn. Het bleek dat ongeveer de helft van de kinderartsen een colloïdale oplossing gaf en de andere helft een kristallijne oplossing. Na het maken van de richtlijn hebben we deze voorkeur opnieuw geëvalueerd en het bleek nu dat het merendeel van de kinderartsen de richtlijn volgde (>90%). We hebben de betrouwbaarheid van dit zelfgerapporteerde gedrag gecontroleerd door het te vergelijken met het in de medische status gedocumenteerde vulmiddel bij een serie kinderen met shock. Dit bleek goed overeen te komen. We denken dat bovengenoemde systematische aanpak geresulteerd heeft in de gemeten verandering in voorkeur en voorschrijfgedrag. De volgende fase van actieve implementatie kon zich dus beperken tot de kleine groep kinderartsen en andere (kinder)subspecialismen die de richtlijn nog niet volgden. Opnieuw werden belemmerende en bevorderende factoren geïdentificeerd en werden de implementatiestrategieën hierop aangepast. Dit resulteerde in een verandering in het voorschrijfgedrag: een verdere toename van het gebruik van fysiologisch zout in plaats van albumine en andere colloiden.

We concluderen dat iedere richtlijn voor het klinisch handelen al in de ontwikkelingsfase aandacht moet besteden aan de implementatie. De doelgroep moet betrokken worden bij het maken van de richtlijn en belemmerende en bevorderende factoren moeten geïdentificeerd worden. Op basis hiervan worden vervolgens de implementatiestrategieën ontwikkeld. Het ontwikkelingsproces moet transparant zijn: er moet een systematisch literatuuroverzicht gemaakt worden en alle andere overwegingen om tot de aanbevelingen te komen moeten expliciet vermeld worden. Tot slot is een toegankelijke, begrijpelijke en attractieve vormgeving van belang met bijvoorbeeld samenvattingen van de adviezen op een plastic kaartje.

Het maken van een goede evidence-based en expliciete richtlijn voor het klinisch handelen is zeer tijdrovend, met name als er een systematisch literatuuroverzicht gemaakt moet worden, en duurt gemiddeld 1,5 jaar. Internationaal zijn er inmiddels vele evidence-based kindergeneeskundige richtlijnen gemaakt en het is dan ook interessant om te weten of die niet ook door ons in Nederland te gebruiken zijn. De vraag is dan of deze richtlijnen kwalitatief goed genoeg zijn en of ze toepasbaar zijn op de Nederlandse kindergeneeskundige situatie.

Hoofdstuk 7 beschrijft een studie waar gezocht is naar bestaande internationale evidence-based kindergeneeskundige richtlijnen. De lijst met gevonden richtlijnen (totaal 215) is naar een groep van 51 kinderartsen gestuurd met de vraag over welke onderwerpen er behoefte was aan een richtlijn. Van de 10 meest genoemde klinische onderwerpen (obstipatie, urineweginfecties, hoofdtrauma <2 jaar en >2 jaar, diabetische ketoacidose, sedatie bij ingrepen, antiemetica bij chemotherapie, vesicoureterale reflux, koorts zonder focus <2 maanden en >2 maanden), zijn alle bestaande richtlijnen (totaal 17) beoordeeld door 3 kinderartsen met behulp van een internationaal geaccepteerd instrument om richtlijnen op kwaliteit te beoordelen: het AGREE

(Appraisal of Guidelines for Research and Evaluation) Instrument. De richtlijnen waren over het algemeen van goede kwaliteit; van de 17 richtlijnen werden 14 door alle drie de kinderartsen als '(sterk) aan te bevelen' voor gebruik in Nederland beoordeeld. Met name het tijdrovende 'evidence' gedeelte hoeft dan niet opnieuw uitgevoerd te worden, daar er al een systematisch literatuuroverzicht gemaakt is. Om het gebruik van deze richtlijnen in de Nederlandse kindergeneeskundige praktijk te vergroten is het wel van belang dat op basis van de 'evidence' eigen aanbevelingen worden geformuleerd door een nationale multidisciplinaire klankbordgroep. Hierbij spelen opnieuw de 'andere overwegingen' een belangrijke rol en die kunnen lokaal verschillen: 'the evidence is global, the recommendations are local'.

In **Hoofdstuk 8** worden op basis van de door ons verrichte studies en de opgedane kennis en ervaring op dit gebied, enkele uitdagingen voor de toekomst besproken om het nog immer bestaande 'gat' te overbruggen tussen klinisch wetenschappelijk onderzoek en de uiteindelijke klinische beslissingen in de praktijk van de kinderarts. De uitdagingen bestaan uit:

1 *Het gebruiken van daadwerkelijk alle bestaande evidence uit afzonderlijke studies om de klinische beslissingen op te baseren.* Om dit doel te bereiken zullen alle afzonderlijke studies rondom een klinische vraagstelling samengevat dienen te worden in systematische literatuuroverzichten. Er zijn er nog vele nodig in de kindergeneeskunde …

2 *Het voorkomen van onnodige duplicatie van inspanningen*. Zoals eerder vermeld, kost het maken van een systematisch literatuuroverzicht en een evidence-based richtlijn veel tijd en vereist het specifieke vaardigheden om de methodologische kwaliteit van de studies te beoordelen. Het is dus van groot belang dat bestaande SRs en richtlijnen gemakkelijk te vinden zijn in elektronische databases door zowel clinici als onderzoekers. Er worden aanbevelingen gedaan over wat er nodig is om SRs en richtlijnen in de toekomst makkelijker te kunnen vinden.

3 *Het verbeteren van de kwaliteit van het klinisch wetenschappelijk onderzoek in de kindergeneeskunde.* Er worden enkele algemene aanbevelingen gedaan om de kwaliteit van primaire studies, systematische literatuuroverzichten en evidence-based richtlijnen te verbeteren.

4 *Het daadwerkelijk vertalen van waardevolle inzichten verkregen uit wetenschappelijk onderzoek naar de klinische praktijk.* Een belangrijk hulpmiddel hierbij zijn evidence-based richtlijnen. Er worden aanbevelingen gedaan hoe toekomstige richtlijnen ontwikkeld en geïmplementeerd dienen te worden.

5 *Het vergroten van de kennis omtrent verschillende onderzoeksmethodes onder de gebruikers ervan*. Om uiteindelijk de resultaten van klinische studies te vertalen naar beslissingen in de praktijk, moeten artsen in staat zijn om studies op methodologische kwaliteit te beoordelen.

Tot slot concluderen we dat de essentiële elementen om het gat tussen onderzoek en de klinische praktijk te dichten zijn 1) **Transparantie** van het gehele proces van studieontwerp tot rapportage van de resultaten – dit geldt voor primaire studies, systematische literatuuroverzichten en richtlijnen, 2) **Methodologische kwaliteit** van primaire studies, systematische literatuuroverzichten en richtlijnen, en 3) **Samenwerking** tussen clinici, richtlijnmakers, onderzoekers, geldschieters en patiënten.

Co-authors

Emma Children's Hospital/Academic Medical Centre, Amsterdam
Department of Pediatric Clinical Epidemiology
Johanna H. van der Lee, MD, PhD; Martin Offringa, MD, PhD
Department of Neonatology
Anne A.M.W. van Kempen, MD, PhD; Joke H. Kok,MD, PhD; Martin Offringa, MD, PhD
Department of Pediatric Intensive Care
Albert P. Bos, MD, PhD
Department of Pediatric Gastroenterology
Merit M. Tabbers, MD

Sophia Children's Hospital/Erasmus Medical Centre, Rotterdam
Department of Pediatrics
Carsten R. Lincke, MD, PhD

Wilhelmina Children's Hospital/University Medical Centre Utrecht, Utrecht
Department of Pediatric Intensive Care
Casper W. Bollen, MD, PhD

Princess Amalia Children's Clinic/Isala Klinieken, Zwolle
Department of Pediatrics
Paul P.L. Brand, MD, PhD

Baylor College of Medicine, Houston, USA
Department of Pediatrics
Virginia A. Moyer, MD, MPH

Alberta Research Centre for Child Health Evidence/University of Alberta, Edmonton, Canada
Department of Pediatrics
Terry P. Klassen, MD, MSc; Lisa Tjosvold, BA, MLIS

United Kingdom Cochrane Centre, Oxford, United Kingdom
Carol Lefebvre, MSc

Dankwoord

Het blijft een raar idee dat het dankwoord de meest gelezen paragraaf van een proefschrift is. Als ik iets wil zeggen, moet ik het dus eigenlijk hier doen. Dat treft, want mijn belangrijkste slotconclusies **'transparantie**, **kwaliteit en samenwerking'** zijn thema's die tussen de regels door ook in mijn dankwoord vaak terug zullen komen. Oprecht kan ik zeggen dat ik met heel veel plezier aan dit proefschrift heb gewerkt. Door de jaren heen ben ik me steeds meer gaan beseffen dat plezier in het werk niet alleen door de inhoud wordt bepaald, maar vooral ook door de mensen met wie je samenwerkt. Ik heb geboft. Daarom wil ik hier dan ook een aantal mensen in het bijzonder bedanken.

First things first, ik begin met mijn 4 prachtige mannen! Liefste **Etienne**, ik weet dat je erg benieuwd bent naar wat ik over jou ga schrijven, gezien je hints de afgelopen tijd naar wat er toch vooral niet(s) moest komen te staan. L, na ruim 20 jaar weet ik dat zo ongeveer wel. Zonder jou was dit proefschrift er (nu) niet gekomen. Het combineren van werk met 3 kleine kinderen is niet altijd makkelijk, zeker niet wanneer je het liefst alles zelf doet. Jij bent de emancipatie voorbij; je vindt het volstrekt normaal dat de taken thuis eerlijk verdeeld worden. Als iemand jou vroeg 'pas jij vandaag op de kinderen' reageerde je verbaasd en zei 'nee, ik zorg gewoon voor mijn kinderen'. Je hebt me alle ruimte gegeven om te doen wat voor mij belangrijk was, zowel privé als op het werk. Heerlijk was het om te weten dat er door jou liefdevol voor de jongens werd gezorgd als ik aan het werk was en daardoor heb ik mijn werkdagen altijd efficiënt kunnen benutten. Al mijn manuscripten heb jij als eerste gelezen en voorzien van zeer helder en bruikbaar commentaar. Dit is slechts één van je vele talenten. Ik zie uit naar het moment dat we weer wat meer tijd hebben om *samen* dingen te kunnen doen en wie weet voor een nieuw avontuur ...
Mick, Vosse en Rowan, ik geniet dagelijks van jullie eerlijkheid, warmte en levensvreugde! Ik heb altijd gezegd dat als ik ooit zou promoveren, er nooit in mijn boekje zou komen te staan dat ik te weinig tijd voor mijn kinderen heb gehad. Het was ook helemaal niet moeilijk, jullie zijn veel te leuk en belangrijk voor me.

Martin Offringa, mijn enige, unieke promotor. Je bent eigenlijk veel meer geweest dan een promotor; je hebt me in allerlei dingen gecoacht. Ik ben denk ik (zeker!) een behoorlijk kritische promovenda geweest, maar wat ik zei was tenminste altijd oprecht. Zo geldt dat ook voor het nu volgende. Ik weet nog goed dat we in het AMC zaten te lunchen en jij mij vroeg om een richtlijn te gaan maken. Dit was 'nieuw en

hot' en je enthousiasme werkte zeer aanstekelijk. Ik was bijna klaar met mijn opleiding tot kinderarts en voelde toen al dat dit een belangrijke stap in mijn carrière kon zijn. Promoveren was lange tijd niet aan de orde, ik moest eerst maar eens wat laten zien. In de daaropvolgende jaren hebben we vele projecten samen gedaan en heb ik verschrikkelijk veel van je geleerd. Je bruist van de energie en nieuwe ideeën. Je hebt me de vrijheid gegeven om de dingen te doen die ik graag wilde doen, me altijd daarin gesteund en je zorgde ervoor dat ik ook daadwerkelijk de tijd kreeg om de projecten tot een goed einde te brengen. Je hebt me in aanraking gebracht met een aantal buitenlandse 'topdogs' en ik heb het voorrecht gehad met hun samen te mogen werken. Ik bewonder je analytisch vermogen, je scherpte en snelle denken. En dat alles ging altijd gepaard met de nodige humor. Ik heb genoten van je prachtige revisies van mijn manuscripten. You're the best! Ik zal het nog missen, tenzij … Heel veel dank voor alles wat je voor me gedaan hebt!

Leden van de promotiecommissie. **Prof. dr. W.M.C. van Aalderen, Prof. dr. P.M.M. Bossuyt, Dr. P.L.P. Brand, Prof. dr. H.N. Caron, Prof. dr. R.P.T.M. Grol, Prof. dr. J.H. Kok, Dr. R.J.P.M. Scholten**, hartelijk dank voor het zitting willen nemen in de promotiecommissie. Ik hoop op een plezierige en interessante discussie de 24^{e} april!

Hugo Heymans, ik wil je bedanken voor het mogelijk maken van mijn toch wel bijzondere aanstelling in het AMC.

Anne van Kempen. Wat is het toch mooi dat de hypoglycemie trial, waarvoor we samen de basis hebben gelegd in hoofdstuk 3, gehonoreerd is door ZonMw. Je hebt er als penvoerder heel veel tijd en bezieling in gestoken. Ook nu vraagt de voorbereiding veel tijd van je. Ik waardeer het zeer dat je me toch met rust hebt willen laten om eerst mijn boekje af te ronden. Ik ga er nu in volle vaart tegen aan en hoop dat we in de komende jaren kunnen laten zien dat Nederlandse kinderartsen goed kunnen samenwerken in dit soort multicenter studies en … dat we eens wat antwoorden gaan genereren.

Merit Tabbers. Als een wervelwind begon jij aan het implementatieproject. Je hebt een fantastisch organisatorisch vermogen en voor mij ben je dé 'implementatoloog'! Dank voor je inzet en de gezellige bijeenkomsten die we hebben gehad.

Medeauteurs van mijn artikelen: **Casper Bollen, Bert Bos, Paul Brand, Anne van Kempen, Joke Kok, Hanneke van der Lee, Carsten Lincke en Merit Tabbers**. Dank voor jullie belangrijke bijdrage aan de verschillende artikelen en de plezierige samenwerking. **Virginia Moyer, Terry Klassen, Carol Lefebvre and Lisa Tjosvold**. It was very inspiring for me to work with you all. Thank you for all the beautiful

work and I am very grateful for your comments on drafts of several chapters of this thesis.

Medische Bibliotheek. **Marjan Loep** heeft veel zorgvuldig werk gestoken in het handmatig zoeken naar kindergeneeskundige systematische literatuuroverzichten in de medische tijdschriften. Zeer spijtig dat ze het resultaat van haar werk nooit heeft kunnen zien. **Arnold Leenders en Marcel van der Paardt** hebben het met verve van haar overgenomen. Dank jullie wel, dat jullie mij alle gegevens op zo'n korte termijn konden aanleveren.

Rob Scholten. De basis voor mijn EBM interesse is eigenlijk gelegd bij het Dutch Cochrane Centre. Ik heb het altijd zeer plezierig gevonden om met je samen te werken. Je bent transparant en levert kwaliteit. Ik zie er dan ook naar uit om samen met het DCC team aan nieuwe, uitdagende projecten te beginnen.

Joke Kok. Jij hebt een belangrijke rol gespeeld bij een aantal van mijn projecten en ik wil je bedanken voor je altijd zeer gedegen commentaar en adviezen. Ook heb jij het mogelijk gemaakt dat ik samen met Anne de hypoglycemie trial ga coördineren. Mijn dank daarvoor en ik zie uit naar de samenwerking!

Wim van Aalderen. Ik wil je bedanken voor het mogelijk maken om naast mijn vele andere projecten toch ook een stukje kliniek te kunnen blijven doen bij de kinderlongziekten. Ik bewonder je altijd goede humeur, rust en flexibiliteit.

KEK- collega's. **Maruschka Merkus**, het was gezellig om samen met jou in de 'kast' op H3 te zitten. Heerlijk om een goede methodoloog binnen handbereik te hebben en dank voor je bereidheid me altijd ad hoc te helpen**. Karin Fijnvandraat**, mijn strategische female coach. Dank voor je interesse en eerlijkheid. Door de tijd heen heb je me een aantal essentiële tips gegeven. **Hanneke van der Lee,** dank voor je zeer heldere commentaar op een aantal hoofdstukken van mijn proefschrift. Ik heb er veel aan gehad. **Anja Brakenhoff**, dank voor de ó zo belangrijke secretariële ondersteuning.

Mijn paranimfen. **Miranda van Turennout**. Lieve Miran, wij zijn al 30 jaar bevriend en hebben vele hoogtepunten gehad (met op stip onze reis door Azië) en eigenlijk geen dieptepunten. Jij bent me al lang geleden voorgegaan in dit 'spelletje' en bent me nog steeds ruim voor met mooie publicaties in Science en Nature. Ondanks je succes ga je nu toch een ander pad inslaan. Good for you! Het gaat je vast lukken. **Henk ter Horst**. Wij zijn tijdens de co-schappen bevriend geraakt. We hebben het heel wat keren licht zien worden … zowel binnen als buiten het ziekenhuis. Samen hebben we onze opleiding tot kinderarts gevolgd in het AMC en in Alkmaar. Je was altijd de ideale collega

voor mij: een goede dokter, collegiaal en veel humor. Geweldig dat jullie mijn paranimfen zijn, we gaan er een mooi feestje van maken!

Erik Faas, goede en geestige vriend. Hartelijk dank voor je adviezen en bemiddeling bij de vormgeving en het drukken van mijn proefschrift.

Annelies Bast. Dank je wel voor de zeer zorgvuldige manier waarop je mijn boekje vorm hebt gegeven. Ik weet het, ik wilde nog wel eens van gedachten veranderen. Gelukkig was dat voor jou nooit een probleem.

Lieve papa en mama, jullie hebben me zoveel (mee)gegeven. Zelfvertrouwen, je talenten benutten, eigen keuzes maken, een positieve levensinstelling en vooral ook genieten van het leven. Mam, het is heerlijk dat ik altijd op je kan rekenen. Ik heb bewondering voor je onbaatzuchtigheid. Pap, ook jij bent er altijd volop voor ons geweest en ik ben je dankbaar voor zoveel mooie herinneringen. Helaas is het voor jou nu een moeilijke tijd. Zoals mijn schoonmoeder zo mooi zegt 'Iedereen wil oud worden, maar niemand wil het zijn'. Jij hebt de laatste tijd wel erg aan den lijve moeten ondervinden dat de ouderdom met gebreken komt. Ik hoop op betere tijden … Dit proefschrift is voor jullie. Mijn lieve zussen, **Michèle en Sacha**. Ik zei het al, ik heb geboft, zo ook met jullie. Dank voor de steun, interesse en gezelligheid.

Curriculum vitae

Nicole Boluyt werd geboren op 14 februari 1967 in Wassenaar. Zij behaalde in 1985 haar eindexamen VWO aan het 'Vlietlandcollege' te Leiden, waarna zij Geneeskunde ging studeren aan de Universiteit van Amsterdam. In 1990 behaalde zij haar doctoraal en in 1993 volgde het artsexamen 'cum laude'.

Vervolgens werkte zij gedurende anderhalf jaar als onderzoeker op de afdeling Klinische Epidemiologie en de afdeling Kinderlongziekten van het UMC Groningen aan een internationaal samenwerkingsproject betreffende 'de effecten van wintersmog op de luchtwegen bij kinderen met astma'.

In 1994/1995 heeft zij gedurende een jaar gewerkt als arts-assistent (Senior House Officer) kindergeneeskunde in verschillende ziekenhuizen in het Verenigd Koninkrijk.

Hierna werd zij arts-assistent kindergeneeskunde in het Medisch Centrum Alkmaar, en startte zij aldaar in 1996 met de opleiding tot kinderarts. Vanaf 1998 volgde zij het academische deel van de opleiding in het Emma Kinderziekenhuis/AMC Amsterdam en in maart 2001 voltooide zij haar opleiding tot kinderarts.

Gedurende de daaropvolgende 2 jaar werkte zij bij het Dutch Cochrane Centre/ AMC en heeft zij tevens projecten op het gebied van evidence-based kindergeneeskunde gedaan die uiteindelijk hebben geresulteerd in dit proefschrift. In 2003 is zij een fellowship klinische epidemiologie begonnen bij de afdeling Klinische Epidemiologie in de Kindergeneeskunde/AMC. Gedurende deze periode was zij tevens als kinderarts werkzaam bij de afdeling Kinderlongziekten, wat zich voornamelijk toespitste op polikliniche werkzaamheden.

Vanaf maart 2007 gaat zij samen met Anne van Kempen een multicenter trial coördineren naar de effecten van een lage bloedsuiker bij pasgeborenen op de latere psychomotore ontwikkeling. Deze studie is het resultaat geweest van een systematisch literatuuroverzicht beschreven in hoofdstuk 3 van dit proefschrift. Daarnaast gaat zij opnieuw bij het Dutch Cochrane Centre werken en zal zij onderwijs geven in evidence-based medicine.

Nicole is getrouwd met Etienne van Banning en ze hebben 3 zoons, Mick (7), Vosse (4) en Rowan (2).