

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI[®]

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

**SYSTEMATIC REVIEWS OF THE EFFECTS OF TWO APPROACHES TO
ENCOURAGE THE USE OF STRICT DIAGNOSTIC CRITERIA FOR THE ONSET
OF ACTIVE LABOUR**

Leeanne Lauzon

**A thesis submitted in conformity with the requirements for the degree of Master of
Science, Graduate Department of Nursing Science
University of Toronto.**

© Copyright by Leeanne Lauzon 1998



National Library
of Canada

Acquisitions and
Bibliographic Services

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale
du Canada

Acquisitions et
services bibliographiques

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-40830-2

**SYSTEMATIC REVIEWS OF THE EFFECTS OF TWO APPROACHES TO
ENCOURAGE THE USE OF STRICT DIAGNOSTIC CRITERIA FOR THE ONSET
OF ACTIVE LABOUR**

Leeanne Lauzon

**A thesis submitted in conformity with the requirements for the degree of Master of
Science, Graduate Department of Nursing Science, University of Toronto**

ABSTRACT

The evidence regarding benefits and risks of health care interventions can facilitate rational, informed decision-making. A systematic review may reveal that a health care intervention has consistent effects, or it may show that treatment effects vary considerably. Applicability of results to populations, settings, and treatment differences are clarified. Results are made more reliable and effects of bias and random error limited through use of an explicit, empirically-based method.

Two systematic reviews were undertaken for the Pregnancy and Childbirth Group of the Cochrane Collaboration. The first determined the effectiveness of caregivers' use of specific criteria for diagnosing labour; the second assessed the effects of teaching pregnant women specific criteria for self-diagnosis of labour onset. Each included one study, involving 209 and 245 women respectively.

Both approaches to promote the accurate diagnosis of labour may be helpful, but further research is required before recommendations for practice can be made.

TO WHOM IT MAY CONCERN:

I, LEEANNE LAUZON, authorize the National Library of Canada to microfilm the following pages/graphs/charts or appendix out of the attached thesis as is: _____, though they may not be legible when reproduced. For reference, a bound copy of the thesis will be available at the University of Toronto's Science & Medicine Library.

L. Lauzon

Signature

12 Aug 98

Date

TABLE OF CONTENTS

CHAPTER 1: THE PROBLEM.	5
Background.	5
Review of the Literature.	7
The Diagnosis of Active Labour.	7
Dystocia.	11
Antenatal Education.	15
Summary.	17
CHAPTER 2: THE REVIEWS.	19
Caregivers' Use of Strict Criteria for the Diagnosis of Active Labour in Term Pregnancy.	20
Cover Sheet.	20
Abstract.	21
Background.	22
Objectives.	23
Materials and Methods.	24
Results.	26
Discussion.	27
Conclusions.	27
Characteristics of Included Studies.	29
References.	30
Table of Comparisons.	32
MetaView.	33
Antenatal Education for the Self-Diagnosis of the Onset of Active Labour in Term Pregnancy.	47
Cover Sheet.	47
Abstract.	48
Background.	49
Objectives.	50
Materials and Methods.	50
Results.	53
Discussion.	53
Conclusions.	54
Characteristics of Included Studies.	55
References.	56
Table of Comparisons.	58
MetaView.	59
CHAPTER 3: CONCLUSIONS.	61

CHAPTER 1

The Problem

Background

A systematic review is a means of identifying effects of health care interventions, and establishing the appropriate application of research results across populations, settings, and differences in treatment (Mulrow and Oxman, 1997). Traditionally used to make large quantities of scientific data more manageable, meta-analyses serve to increase the power of the pooled results of smaller studies and of those in which a treatment effect was not statistically significant (Mulrow, 1994). Studies that explore the effects of utilising strict diagnostic criteria for the onset of active labour may be so categorized. The mistaken diagnosis of active labour is associated with the diagnosis of dystocia and a resultant cascade of potentially harmful medical interventions. However, no systematic review has been conducted of the evidence supporting the use of diagnostic criteria for active labour. Cervical dilatation, uterine contractions, and the status of the amniotic membranes are criteria typically considered by clinicians for labour diagnosis.

One method of conducting a systematic review is that which is used by the Cochrane Collaboration. The mission of the Cochrane Collaboration is to generate the highest quality systematic reviews of the evidence concerning benefits and risks of health care practices. The method by which these reviews are completed is outlined in Appendix A.

There are two approaches that may be considered when examining the effects of using strict diagnostic criteria for active labour. Obstetrical caregivers may apply specific

criteria to women presenting for labour assessment. Alternatively, women themselves may be taught a specific antenatal education program which would assist them to recognise active labour. Because there are these two complementary approaches to labour diagnosis, two related systematic reviews were completed for the Cochrane Collaboration. One review is titled, "Caregivers' use of strict criteria for the diagnosis of active labour in term pregnancy;" the second is titled "Antenatal education for self-diagnosis of the onset of active labour in term pregnancy."

Review of the Literature

The review of the literature is divided into the following sections: the diagnosis of active labour, dystocia, and antenatal education.

The Diagnosis of Active Labour.

Cardozo and Studd (1985) described the onset of labour as a process rather than an event. Within this process is the transition from early or latent labour to active, progressive labour. There exists no consensus, however, on the specific criteria indicative of this transition. Regardless of the associated difficulties and uncertainty about the onset of active labour, the diagnosis of active progressive labour is a fundamental component of labour care as there are potential repercussions for women who have been incorrectly diagnosed, which may include being diagnosed with labour dystocia (Baskett, 1991; Friedman, 1989; and O'Driscoll, Foley, and MacDonald, 1984). Treating labour dystocia leads to further repercussions for women, as it is associated with an increased incidence of analgesia use and uterine hyperstimulation (Thornton and Lilford, 1994), and an increased Caesarean section rate for fetal distress (Fraser, Krauss, Brisson-Carrol, Thornton, and Breart, 1995).

Several authors (Baskett, 1991; Crowther, Enkin, Keirse and Brown, 1989; Hemminki and Sumukka, 1986; O'Driscoll, *et al*, 1993) emphasized the importance of establishing a correct diagnosis of labour to support the provision of appropriate care and the suitability of intervention planning. Baskett (1991) referred to the practical problem of diagnosing true labour when assessing a number of criteria. While well-advanced labour may be fairly obvious to caregivers and to expectant women, distinguishing true labour in

its early or latent phase from false labour is more challenging and clinically problematic. This may be attributed to the similarity of symptoms (Friedman, 1978), poorly differentiated and apparent overlap of arbitrarily assigned “stages” of labour (Rosen, 1990), and the lack of consensus amongst caregivers about diagnostic criteria.

False labour contractions may be regular and strong, even quite painful, and sometimes may last for hours (Simkin, 1989). False labour may be distinguished from latent labour retrospectively because the contractions do not increase in duration, strength, and frequency, cervical dilatation does not occur, and labour does not become established in the next 24 hours (Arulkumaran, Michelsen, Ingemarsson and Ratnam, 1987; Quinn, Murphy and Gallagher, 1984; Simkin, 1989; Tay, 1991).

There are four principal criteria taken into consideration by clinicians when deciding upon a diagnosis of active labour. These include the character of uterine contractions, (Baskett, 1991; Crowther, Enkin, Keirse, and Brown, 1989; Cunningham, MacDonald, and Gant, 1989; O’Driscoll, Meagher, and Brown, 1993; Oxorn, 1986), the integrity of the amniotic membranes (Baskett, 1991; Crowther *et al.*, 1989; O’Driscoll *et al.*, 1993), the presence of a vaginal “show”(Baskett, 1991; Crowther *et al.*, 1989; O’Driscoll *et al.*, 1993), and cervical dilatation and/or effacement (Cardozo and Studd, 1985; Crowther *et al.*, 1989; Friedman, 1978; O’Driscoll *et al.*, 1993; Studd, 1973; Oxorn, 1986; Peisner and Rosen, 1985). As these criteria are interrelated, they must be considered in relation to one another when considering a diagnosis of labour.

Painful uterine contractions are the key component of active labour, common to virtually all women in labour (Baskett, 1991; O'Driscoll *et al.*, 1993; Oxorn, 1986). While contractions are assessed in terms of frequency, strength and duration, there do not appear to be distinct differences between those of “true” or active labour, and Braxton-Hicks contractions. Those authors who have attempted to distinguish between the contractions of active and “false” labour appear to have offered subjective interpretations of their experiences, as they do not provide references to empirical evidence (Cunningham, MacDonald, and Gant, 1989; Oxorn, 1986). Objective attempts to quantify uterine activity abound, although none has been demonstrated to improve the outcome of labour (Crowther, Enkin, Keirse, and Brown, 1989). There is a lack of evidence that these attempts at quantification enable caregivers to definitively diagnose active labour, quite possibly because they are intended to describe uterine activity rather than to be used as diagnostic tools. Thus, there is a paucity of information available from clinical trials regarding the frequency, duration, and strength of contractions of active labour versus those of latent stage or false labour.

The value of interpreting the status of the amniotic membranes and the presence of vaginal “show” as indicative of active labour is also questionable. The amniotic membranes can rupture at any time during the course of pregnancy or labour, and are therefore not a reliable or sensitive indicator of labour. A “show” of blood-streaked mucus from the vagina is generally considered to be a positive sign that the onset of labour is impending (Cunningham, MacDonald, and Gant, 1989; O'Driscoll, Meagher, and Boylan, 1993). However, cervical changes may begin to occur as early as four weeks prior to the

onset of labour (Hendricks, Brenner, and Krams, 1970), and a rectal or vaginal examination of the cervix performed in the preceding 48 hours may be an irritant to the cervix, resulting in a blood-streaked mucous discharge (Cunningham et al.).

The traditional use of these criteria and the variety of combinations, therefore, appears to have been perpetuated through expert opinion, arbitrary consensus and anecdotal evidence. While Baskett (1991), Crowther, Enkin, Keirse and Brown (1989), Hemminki and Sumukka (1986), and O'Driscoll *et al* (1993) emphasized the importance of correct labour diagnosis, none has empirically tested their proposed diagnostic criteria.

Cervical dilatation is the sole criterion for labour diagnosis that researchers have attempted to subject to scientific scrutiny. Perhaps the most widely cited research in evaluating cervical dilatation in labour is that conducted by Friedman in the 1950's. According to Friedman's initial retrospective chart review of 100 women, latent labour, which is of a variable duration, lasts from 0 cm to 3 cm dilatation of the cervix, and active labour is diagnosed when the cervix has achieved a dilatation of 3 or more centimetres (Friedman, 1967). Despite the limitations of his research, his work is cited by a number of authors (Baskett, 1991; Oxorn, 1986; Peisner and Rosen, 1986) as defining the phases of normal labour, in differentiating between labour that is latent, active, and that which is false. In 1970, however, Hendricks, Brenner, and Kraus found in a descriptive study of 303 women that cervical dilatation may occur in the 4 weeks preceding the onset of active labour, and that within the 3 days before labour the mean dilatation was 1.8 cm and 2.2 cm in primigravidae and multigravidae respectively (Cardozo and Studd, 1985). These

findings, while also limited, render questionable the labour stages defined by Friedman, who acknowledged in later works (1978) that the diagnosis of true progressive labour can only be made accurately in retrospect.

The sensitivities, specificities, and predictive values of the diagnostic criteria for labour, alone or in combination, have not been well-studied or reported. The paucity of high-quality research makes it difficult to employ standards or degrees at which these criteria become significant indicators of active labour (Crowther *et al.*, 1989). A range of criteria that is too narrow, while increasing specificity, has the effect of decreasing sensitivity (Thornton *et al.*, 1994). Without having a solid objective basis for determining active labour, setting arbitrary limits may have the possible effects of reducing the number of women erroneously diagnosed to be in active labour, while simultaneously increasing the number of women in whom true labour dystocia has been overlooked. Conversely, a more broad range of diagnostic criteria exposes a greater number of women to those consequences associated with erroneous labour diagnosis. This apparent lack of evidence to accurately define active labour illustrates the need for a systematic review of the relevant literature.

Dystocia

The term “dystocia” is intended to describe abnormal or non-progressive labour (Baskett, 1991), and is the leading cause of primary Caesarean section (Porreco and Thorp, 1996). Failure of the cervix to dilate within a prescribed period of time, as would be expected in active progressive labour, results in a diagnosis of dystocia (Baskett, 1991; Friedman, 1989; and O’Driscoll, Foley, and MacDonald, 1984). This diagnosis, therefore,

is directly associated with one's understanding of normal labour, which includes by necessity the criteria one applies to determine that active labour has commenced. As it is only in active labour that cervical effacement and dilatation would be evident, it is inaccurate and inappropriate to diagnose dystocia in latent or false labour (Rosen, 1990).

Dystocia is typically treated with an approach to correct inefficient uterine contractions. Oxytocin augmentation and artificial rupture of the amniotic membranes (amniotomy) are most commonly utilised in the attempt to correct contractions that are not strong enough, frequent enough, or of insufficient duration (O'Driscoll *et al*, 1993; Thornton *et al*, 1994). As these methods are also commonly used to induce labour, an initial questionable diagnosis of active labour may become inconsequential. However, there are potential consequences associated with the use of oxytocin augmentation and early amniotomy (Thornton *et al*, 1994). Oxytocin augmentation may increase the incidence of analgesia use and uterine hyperstimulation (Thornton *et al*, 1994.). Uterine hyperstimulation may result in decreased placental perfusion, fetal hypoxia, or uterine rupture. Early amniotomy is associated with an increased Caesarean section rate for fetal distress (Fraser, Krauss, Brisson-Carrol, Thornton, and Breart, 1995).

Stewart, Dulberg, Arnill, Elmslie, and Hall (1990) noted in a retrospective chart review of 3887 women that 41% of women who had undergone Caesarean section for dystocia had not yet achieved active labour, and thus the diagnosis of dystocia was apparently incorrect. This rate may be somewhat elevated due to the study's diagnostic criteria for active labour of 4 cm dilatation, determined from the average of 3 to 5 cm from

the labour curves of individual participants. Had active labour been diagnosed in accordance with the criteria of 3 cm dilatation suggested by O'Driscoll *et al.* (1969), Friedman (1967), and the Canadian Consensus Panel on Cesarean Birth (1985), the resultant percentage might have been somewhat less.

There are inherent difficulties associated with a retrospective study design. The study by Stewart *et al.* (1990) illustrates the problem of defining what is meant by a diagnosis of dystocia when there exists no consensus on the definition. In order to capture all possible diagnoses of dystocia, the researchers included all women who had on their hospital documentation records any of twenty-four terms that may have indicated dystocia. Despite using one specially trained nurse for data extraction, this broad method of inclusion may have influenced their findings through increasing, perhaps inaccurately, the numbers of women experiencing Caesarean section for apparent dystocia. Additionally, the investigators did not seek to assess the validity of the chart information, as they were attempting to determine physician decision-influencing factors in the context of each situation. Despite the limitations associated with this retrospective design, the design has the advantage of not influencing the behaviour or decision-making processes of the physicians. A Caesarean section rate for dystocia, without true labour, of 41 % as described by Stewart *et al.* (1990) is related to the use of a broad range of diagnostic criteria for active labour, so broad as to significantly decrease the specificity of these criteria.

In 1986, Peisner and Rosen determined that 90 % of their descriptive study sample of 1699 women had reached active progressive of labour by 5 cm dilatation. Confounding

this finding is their decision to use Friedman's (1982) parameters for defining active labour to use cervical dilatation as the sole diagnostic criteria for active labour, and to eliminate all women diagnosed with labour dystocia. When women who developed labour dystocia after having achieved 5 cm dilatation were included, the authors found that 74% of participants were in active labour by 5 cm dilatation. This finding supports the parameters set for active labour diagnosis by Stewart *et al.* (1990), whose results serve to illustrate an important potential resultant effect of failing to utilise empirically-based strict diagnostic criteria for active labour.

Dystocia is the principal indication for primary Caesarean section, particularly in nulliparous patients, and is therefore the largest contributing factor to repeat Caesarean section in subsequent pregnancies (Baskett, 1991; Porreco and Thorp, 1996). As nulliparas are more likely to be affected by such a diagnosis, the criteria by which nulliparas are determined to be in active labour and progressing "normally" must be called into question. An augmentation rate of 55% for nulliparas (as reported by O'Driscoll *et al.*, 1973) suggests that the majority of nulliparas progress in labour more slowly than what is considered to be normal (Studd, 1973). This statement is supported by findings of The Canadian Early Amniotomy Study Group (Fraser, Marcoux, Moutquin, and Christen, 1993). In their randomised controlled trial of the effect of early amniotomy on the risk of dystocia in nulliparous women, dystocia was reported at rates of 34 % and 45 % in the study and control groups respectively. Thus, 45 % of the women being conservatively managed received oxytocin augmentation. With such a relatively high labour

augmentation rate amongst this control group, it would seem that the perception of what is optimal normal labour progress has replaced that which is an actual normal labour progress, probably in relation to a narrowing of the criteria by which normal labour progress is defined.

The parameters for normal progression of labour as defined by O'Driscoll *et al.* (1973) appear to be too narrow, as evidenced by their high rate of oxytocin augmentation. Both the findings of Fraser *et al.* (1993) and those of Stewart *et al.* (1990) appear to result from a lack of clear empirically-based parameters by which to diagnose and monitor active progressive labour. These studies illustrate the need for specific research-based diagnostic criteria for the onset and progress of active labour.

Antenatal Education.

In view of the potential consequences associated with an erroneous diagnosis of active labour, it would seem to be advantageous to seek objective, empirically sound means to accurately diagnose active labour. When considering the use of strict diagnostic criteria for labour as a means to reduce the incidence of mistakenly diagnosing active labour, two approaches may be taken into account. First, obstetrical caregivers may be taught to apply specific criteria to women presenting for labour assessment. Alternatively, women themselves may be taught a specific antenatal education program which would assist them in recognising signs of active labour, and to help them make decisions about when to be assessed for labour progress, and support if required.

Education is a process of acquiring knowledge and skills, so that one may develop understanding, attitudes, and values. There are numerous studies of education as an

intervention, the discussion of which is beyond the scope of this thesis. As concerns antenatal education, it would seem quite reasonable to assume that pregnant women would be receptive to information and advice related to pregnancy and childbirth.

There are numerous educational resources available to pregnant women and their partners and families. Books and periodicals, structured prenatal education classes, as well as family, health care, and social contacts are all sources of information to the expectant woman, each resource encompassing its own underlying goals, purposes and ideologies. Thus, it is difficult to determine the effects on women of antenatal education in general as a distinct intervention (Simkin and Enkin, 1989). Physicians, midwives, and antenatal educators are likely to be the most accessible resources for obtaining and providing specific information about labour self-assessment, yet there is little evidence to indicate whether teaching women how to assess their labours actually makes a difference in the outcome of their pregnancies.

Kramer (1996) conducted a systematic review of the randomised controlled trials of the effects of giving women nutritional advice in pregnancy on pregnancy outcome. Four studies, involving a total of 1098 women, were included in this review, although Kramer notes that the methodological quality of these studies is not high. Twelve outcomes were measured, relating to dietary intake, and maternal and neonatal outcomes. Kramer (1996) concluded that advising pregnant women to increase their energy and protein intake resulted in a significant albeit modest increase in their intake of these nutrients. These findings illustrate the utilization of a specific antenatal education

intervention to create a desired effect. Similarly it is hypothesized that a specific curriculum intended to facilitate self-diagnosis of labour may have the effect of reducing the numbers of women who present for labour care when labour is not yet established. This, in turn, may reduce the occurrence of mistaken diagnosis of active labour and its associated consequences. It is hypothesized, then, that a structured program educating women on recognising the signs of labour may be more beneficial than the usual sources of information and advice. This hypothesis was investigated in a Cochrane Review.

Summary

An examination of the relevant literature has illustrated the lack of adequately controlled studies to diagnose and/or define active labour and the resultant potential consequences which included the erroneous diagnosis of dystocia. The diagnosis of dystocia was also noted to lead to cascading medical interventions which are not without consequences. This review has also suggested that women are receptive to instruction in the antenatal period. It was therefore hypothesized that women would also be receptive to information to help them determine true labour. The paucity of high-quality research supporting the application of specific diagnostic criteria for diagnosing active labour as they are currently used (i.e. uterine contractions, status of amniotic membranes, presence of vaginal “show”, and cervical dilatation and effacement) presents a situation worthy of scientific evaluation. It is hypothesized, then, that the adoption of specific diagnostic criteria for active labour may benefit pregnant women through a reduction of mistaken labour diagnoses, or alternatively the diagnosis of dystocia for women who were not truly in active labour, and subsequent reduction of the associated treatments and their potential

consequences. In the attempt to identify specific diagnostic criteria for the onset of active labour, and to support the research hypothesis, two systematic reviews of the evidence were done, following procedures described in the Cochrane Collaboration Handbook Version 4.0 (1997) (see appendix A).

CHAPTER 2

The Reviews

I undertook two structured reviews under the auspices of the Cochrane Pregnancy and Childbirth Group. The method by which these reviews were completed has been developed by The Cochrane Collaboration; it is outlined in Appendix A. The protocols for these reviews have been published in The Cochrane Library (1998, Issue 1).

After registering the review titles with the Collaborative Review Group's coordinator, Sonja Henderson, in May of 1997, an initial draft of the protocols was submitted for review on October 2, 1997. Comments and suggestions from the editor, Prof. G. J. Hofmeyr, and two referees, Drs. D. Jewell and G. Young (received November 10 and 15, 1997) helped to ensure the clarity, succinctness, and thoroughness of the protocols and reviews. Editorial comments were taken into consideration and the protocols accordingly adjusted, and were accepted for publication December 1, 1997. The reviews were submitted for publication in May, 1998, and the texts are included in this chapter.

Caregivers' Use of Strict Criteria for the Diagnosis of Active Labour in Term Pregnancy

Abstract

Objective

To assess the effectiveness of the use by caregivers of specific criteria for diagnosis of active labour in term pregnancy.

Search Strategy

The register of clinical trials maintained and updated by the Cochrane Pregnancy and Childbirth Group and the Cochrane Controlled Trials Register.

Selection Criteria

All randomised controlled trials comparing caregivers' application of strict diagnostic criteria for active labour versus routine practice were considered.

Data Collection and Analysis

Performed by the authors from published articles obtained from trial authors.

Main Results

Women who had experienced Early Labour Assessment were less likely to receive intrapartum oxytocics, and reported higher levels of control during labour and birth.

Conclusions

There may be positive outcomes associated with the implementation of specific diagnostic criteria for active labour diagnosis. A multi-centre RCT would be most helpful to determine the full effect of the use by caregivers of strict diagnostic criteria for active labour.

Background

Timely diagnosis of progressive labour is problematic for caregivers and expectant women. Methods of distinguishing active or progressive labour from latent phase or Braxton-Hicks contractions vary greatly, are not universally applied, and have not been subjected to scientific scrutiny (Thornton and Lilford, 1994). While the diagnosis of active labour is self-evident in retrospect, retrospective diagnosis is of little value in the clinical setting when decisions regarding admission for care and/or intrapartum care interventions are occurring in a relatively brief period of time (Crowther, Enkin, Keirse, and Brown, 1989).

While Braxton-Hicks contractions and even prolonged latent labour may pose no problem for a woman and her fetus, the mistaken diagnosis of active progressive labour is not without consequence. Failure of the cervix to dilate within a prescribed period of time, as would be expected in the active phase of active progressive labour, results in the diagnosis of dystocia (Baskett, 1991; Friedman, 1989; O'Driscoll, Foley, and MacDonald, 1984). One Canadian study found that over 40% of caesarean sections performed for dystocia were done before the establishment of true labour (Stewart, Dulberg, Arnill, Elmslie, and Hall, 1990). Labour that is non-progressive may be augmented with oxytocics in an attempt to correct the inefficiency of the uterus (O'Driscoll et al., 1984; Thornton and Lilford, 1994), which may increase the incidence of analgesia use and uterine hyperstimulation (Thornton and Lilford). Mothers' confidence in their caregivers may be undermined, their self-confidence eroded, and the perception of their birth

experience negatively affected when an incorrect labour diagnosis is amended (O'Driscoll et al., 1984, Simkin, 1996).

Three of the components of the Active Management of Labour protocol (O'Driscoll et al., 1984) have been the subjects of Cochrane Reviews ["Support from caregivers during childbirth", "Amniotomy to shorten spontaneous labour", and "Early oxytocin to shorten labour"]. The question of whether the accurate diagnosis of active labour has been adequately evaluated, however, remains unresolved. The hypothesis of this review is that application of specific criteria for diagnosing active labour benefits those women presenting for labour assessment. Delayed admission to a labour ward for women not yet in active labour may result in a reduced incidence of the above-mentioned consequences for those erroneously admitted for active labour care. The aim of this review is to determine if there is evidence to support the adoption and universal application by caregivers of such criteria. A related review focuses on the effectiveness of teaching pregnant women a set of criteria for self-diagnosis of active labour.

Objectives

To assess the effectiveness of the use by caregivers of specific criteria for diagnosis of active labour in term pregnancy. The main outcomes of interest are:

- rates of Caesarean section for dystocia,
- rates of other intrapartum complications,
- use of oxytocics, analgesics, and other intrapartum interventions,
- mothers' evaluations of their birth experiences and care provided,
- rates of hospital discharge diagnoses of "not in labour" or "false labour",
- rates of out-of-hospital emergencies (e.g. unplanned out-of-hospital births),
- admission rates to special care baby unit/neonatal intensive care unit.

Materials and Methods

Criteria for Considering Studies for This Review

Types of studies.

Randomised controlled trials comparing caregivers' application of strict diagnostic criteria for active labour versus routine practice; violations of allocated management not sufficient to materially affect outcomes; missing data insufficient to materially affect the comparison.

Types of participants.

All pregnant women at term gestation.

Types of interventions.

Application by caregivers of specific criteria for diagnosing active labour in an obstetrical assessment setting, in order to help decision-making regarding the provision of labour care.

Types of outcome measures.

The main outcomes of interest were: rates of Caesarean section; rates of other intrapartum interventions; intrapartum complications; labour augmentation rates; degree of patient satisfaction; rates of hospital discharge diagnoses such as "not in labour" and "false labour;" rates of out-of-hospital emergencies; and neonatal outcomes, including admission rates to neonatal intensive care.

Search Strategy for Identification of Studies

This review has drawn on the search strategy developed for the Pregnancy and Childbirth Group as a whole.

Relevant trials were identified in the Group's Specialised Register of Controlled Trials and the Cochrane Controlled Trials Register. A manual search of the Group's identified resources was conducted for the year immediately preceding the date of the most recent substantive amendment in the attempt to capture all relevant materials. See Review Group's details for more information.

Combinations of search terms used in the manual search and that of the Cochrane Controlled Trials Register for this review included the following:

- labour and education
- labour and recognise
- labour and onset
- labour and active
- triage
- labour and diagnosis
- labour and false
- labour and caregivers
- labour and nurse
- birth and education
- labour or pregnancy
- childbirth

Methods of the Review

The reviewers independently selected and assessed the single trial resulting from the search. Names of authors, related institutions, journals of publication, and study results were known by the reviewers when inclusion criteria were applied. Trials under consideration were evaluated for methodological quality and appropriateness for inclusion,

regardless of results, using standard Cochrane criteria. No identified trials were excluded from this review. Included trial data were processed as described in: Mulrow CD, Oxman AD (eds.), Cochrane Collaboration Handbook [updated 1 March 1997]. In: The Cochrane Library [database on disk and CD-ROM]. The Cochrane Collaboration. Oxford: Update Software; 1996-. Updated quarterly.

Description of Studies

See Characteristics of Included Studies.

Only one trial met the inclusion criteria. The single trial included in this review compared an Early Labour Assessment Program with standard care for the diagnosis of active labour.

Methodological Quality of Included Studies

The single included trial (McNiven 1998) is of excellent quality. Randomisation methods were clear and adequately controlled. Only one randomised patient was lost to follow-up. There is some performance bias noted as some controls (16.3%) were discharged undelivered following direct admission, in comparison to 18.6% of the experimental group. This, however, may be attributable to the routine care to which the control group was randomised.

Results

Women who had experienced Early Labour Assessment were less likely to receive intrapartum oxytocics (OR = 0.44; 95% C.I. 0.24, 0.80) and analgesia (OR = 0.31; CI.

1.26, 7.13), and reported higher levels of control during labour and birth. No other statistically significant results were found.

Discussion

The one trial in this review demonstrated some positive effects of implementing the use by caregivers of specific criteria for diagnosis of active labour in term pregnancy. The use of strict criteria for the diagnosis of active labour may prevent the misdiagnosis of dystocia in latent phase labour and thereby prevent unnecessary, and potentially risky, interventions, including Caesarean delivery. Limitations of this study include the small sample size and the unavoidable fact that participants and their caregivers were unblinded to their study group assignment (they either received initial care in an early labour assessment area or were sent directly to the labour ward). The trial had insufficient power to test the effects of the intervention on rates of Caesarean delivery, unplanned out-of-hospital birth, or other important maternal and neonatal outcomes. A larger multi-centre trial would help to determine the benefits and risks of this intervention.

Conclusions

Implications for Practice

Because only 209 women in a single Canadian centre have been studied, there are no implications for changes to existing practices.

Implications for Research

Although some positive outcomes were associated with the use by caregivers of specific diagnostic criteria for active labour, a multi-centre randomised controlled trial is necessary to determine the risks and benefits of the same or a similar intervention.

Participants should be of sufficient number to allow for the determination of clinically important outcomes, such as the effect on Caesarean section rates, neonatal well-being, rates of intrapartum interventions, and women's evaluations of their care. Potential risks should also be evaluated. These include unplanned out-of-hospital births and the potentially harmful effects of withholding caregiver support and attention to women in early or latent phase labour.

Characteristics of Included Studies

Caregivers' use of strict criteria for the diagnosis of active labour in term pregnancy

Study	Method	Participants	Interventions	Outcomes	Notes
McNiven 1996	Randomised controlled trial, random allocation by sealed, opaque envelopes. Participants were recruited upon presentation to hospital for labour assessment, by labour and delivery nurses trained in the study method. Efforts to eliminate selection bias and detection bias were adequate, and attrition bias was negligible as only one participant was lost to follow-up. Some performance bias is noted as some controls (16.3%) were discharged undelivered following direct admission, in comparison to 18.6% of the experimental group.	209 low-risk nulliparous women at a teaching hospital in Ontario, Canada, were randomly allocated to experimental (n=105) or control (n=104) groups. All were 37+ weeks singleton gestation, with a spontaneous onset of labour. No significant differences in groups with regard to time of recruitment, age, marital status, type of hospital coverage, location of residence, physician, or newborn characteristics. One subject was excluded after randomisation because she was <37 weeks gestation.	Women in the experimental group received Early Labour Assessment, which included the determination of fetal heart rate, maternal blood pressure and urine tests, frequency and duration of contractions, status of amniotic membranes and presence of bloody show. A vaginal examination was performed by a medical intern or an assessment area nurse, and a participant was determined to be in active labour with a cervical dilatation of 3+ cms in the presence of regular, painful, contractions. Participants not found to be in active labour were given advice, support, and encouragement, and were instructed to walk outside or return home (18.6% were sent home). Instruction was given regarding when to return to hospital. Control group participants were admitted directly to the labour and delivery unit.	Caesarean delivery, amniotomy, anaesthesia/analgesia, episiotomy, forceps delivery, vacuum delivery, length of labour, length of stay in labour and delivery and in postpartum unit, oxytocin augmentation, one- and five-minute Apgar scores, participant satisfaction.	Strict adherence to study protocols not followed for all subjects as 16.3% of controls were sent home after assessment in the labour ward.

Doctoral dissertation as basis for McNiven 1998 publication.
RCT = randomised controlled trial

References to Studies Included in This Review

McNiven P. Feasibility study for a multi-centre randomized trial of the effect of an early labour assessment program on the rate of Cesarean delivery. 1996 University of Toronto (unpublished doctoral dissertation).

McNiven PS, Williams, JI, Hodnett E, Kaufman K, and Hannah ME. An early labour assessment program: A randomised, controlled trial. *Birth* 1998; 25(1); 5-10.

Additional References

- Baskett TF (1991). Essential Management of Obstetrical Emergencies (2nd ed).**
Clinical Press Ltd.: Bristol, UK.
- Crowther C, Enkin MW, Keirse MJNC, and Brown I (1989). Monitoring the progress of labour.** In Chalmers I, Enkin MW, and Keirse MJNC (eds.), *Effective Care in Pregnancy and Childbirth* (v.2). Oxford University Press.
- Friedman E (1989). Normal and dysfunctional labour.** In Cohen WR, Acker DB, and Friedman EA (eds.) *Management of Labour* (2nd ed.). Aspen Publishers: Rockville, MD.
- O'Driscoll K, Foley M, and MacDonald D (1984). Active management of labor as an alternative to cesarean section for dystocia.** *Obstetrics and Gynecology*; 63(4): 485-490.
- Simkin P (1996). The experience of maternity in a woman's life.** *JOGNN*; 25(3); 247-252.
- Stewart P, Dulberg C, Arnill A, Elmslie T, and Hall P (1990). Diagnosis of dystocia and management with cesarean section among primiparous women in Ottawa-Carleton.** *Canadian Medical Association Journal*; 142(5): 459-463.
- Thornton JG and Lilford RJ (1994). Active management of labour: Current knowledge and research issues.** *British Medical Journal*; 309: 366-9.

Table of Comparisons

01.00.00 Early labour assessment versus standard care

01.01.00	discharged undelivered
01.02.00	out-of-hospital emergencies
01.03.00	artificial rupture of membranes
01.04.00	intrapartum oxytocics
01.05.00	any intrapartum analgesia
01.06.00	epidural analgesia
01.07.00	intrapartum narcotic/inhalation analgesia
01.08.00	forceps/vacuum extraction
01.09.00	Caesarean section rates (overall)
01.10.00	Caesarean section rates for labour dystocia
01.11.00	perceived control
01.12.00	1-minute Apgar <7
01.13.00	5-minute Apgar <7
01.14.00	neonatal resuscitation
01.15.00	admission to neonatal intensive care

Review: Diagnosis of active labour in term pregnancy

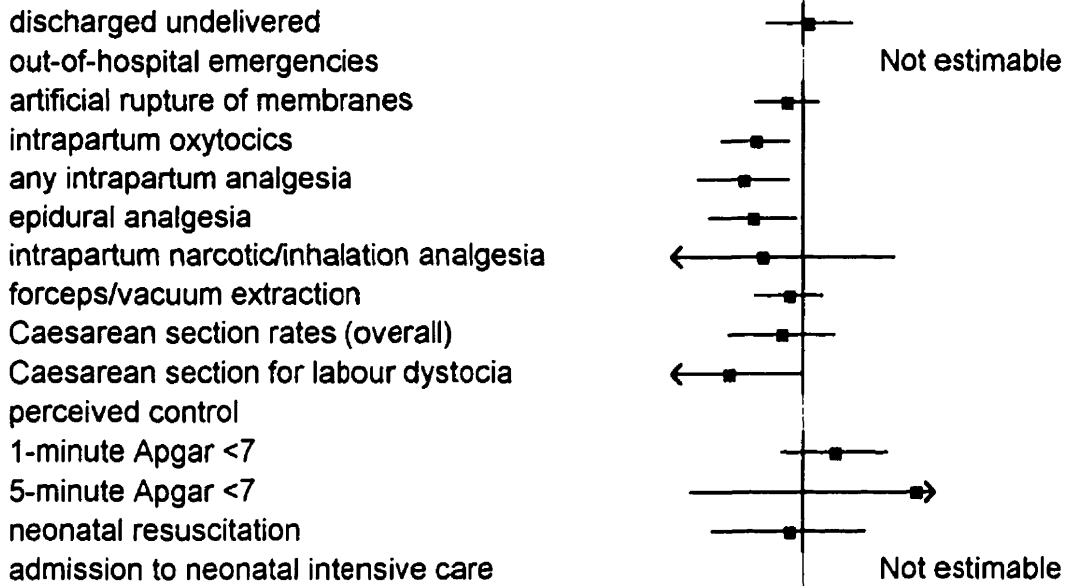
Comparison or Outcome

Peto Odds Ratio

WMD (95%CI)

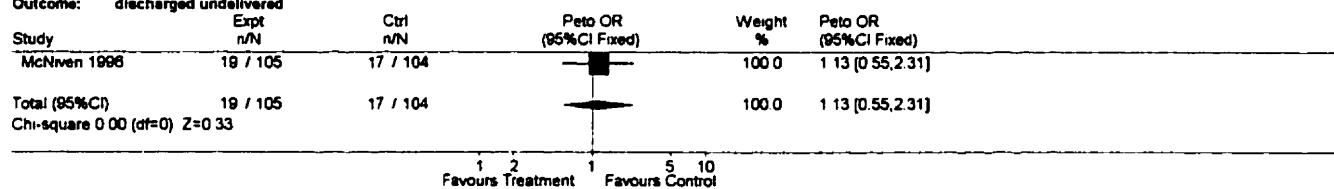
Early labour assessment versus standard care

(95%CI)



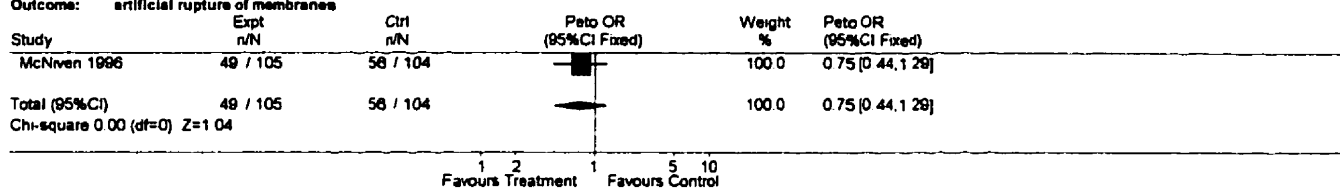
.1 .2 1 5 10 -10 -5 0 5 10

Review: Diagnosis of active labour in term pregnancy
 Comparison: Early labour assessment versus standard care
 Outcome: discharged undelivered



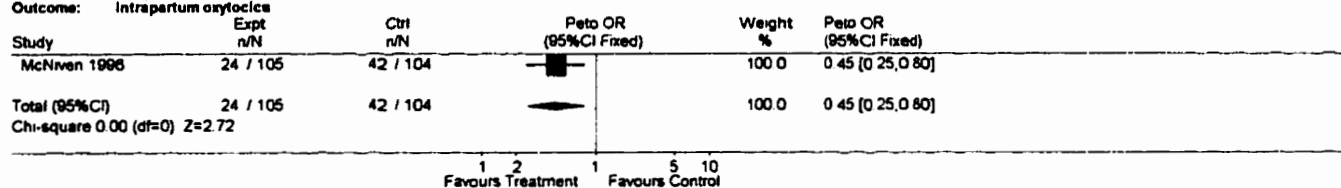
Review: Diagnosis of active labour in term pregnancy
 Comparison: Early labour assessment versus standard care
 Outcome: artificial rupture of membranes

35

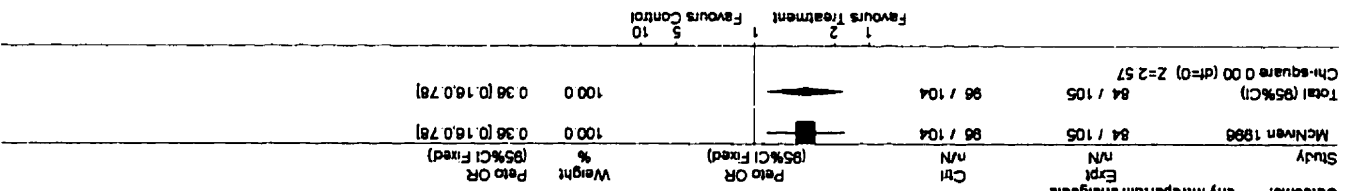


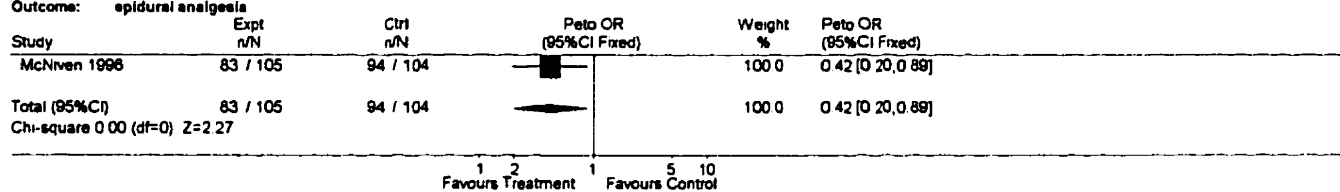
Review: Diagnosis of active labour in term pregnancy
 Comparison: Early labour assessment versus standard care
 Outcome: Intrapartum oxytocics

36



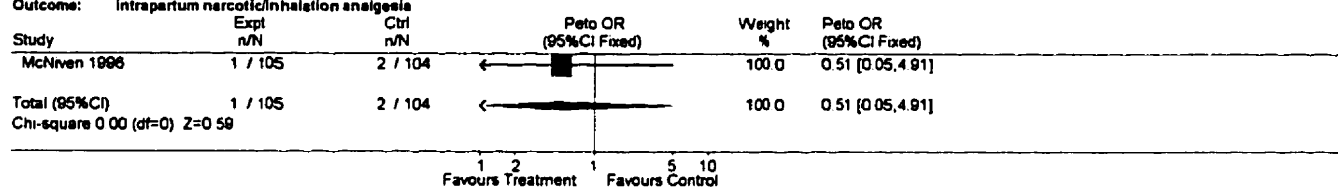
Review: Diagnosis of active labour in term pregnancy
 Comparison: Early labour assessment versus standard care
 Outcome: any intrapartum anaesthesia



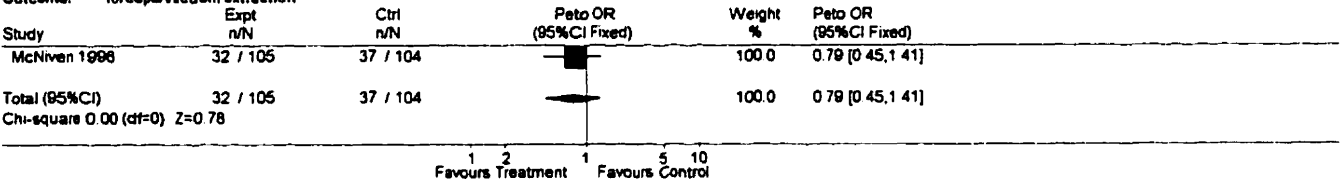


Review: Diagnosis of active labour in term pregnancy
 Comparison: Early labour assessment versus standard care
 Outcome: Intrapartum narcotic/inhalation analgesia

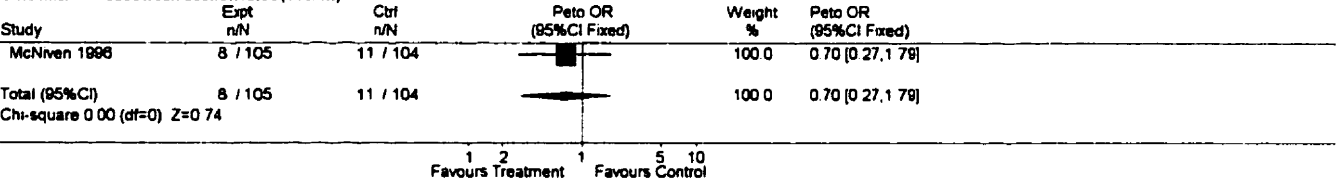
59



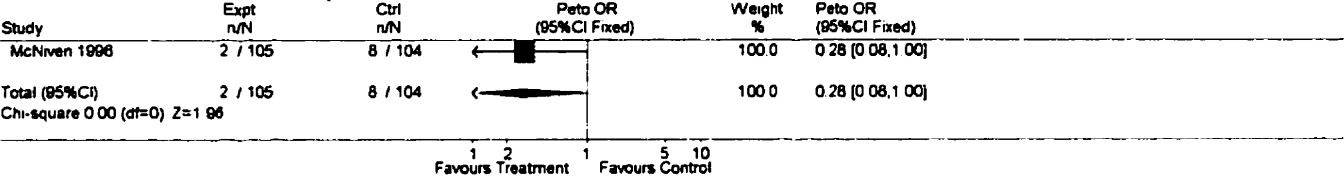
Review: Diagnosis of active labour in term pregnancy
Comparison: Early labour assessment versus standard care
Outcome: forceps/vacuum extraction



Review: Diagnosis of active labour in term pregnancy
 Comparison: Early labour assessment versus standard care
 Outcome: Caesarean section rates (overall)

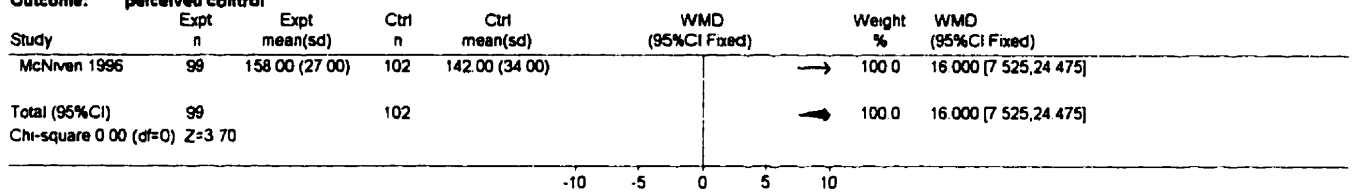


Review: Diagnosis of active labour in term pregnancy
Comparison: Early labour assessment versus standard care
Outcome: Caesarean section for labour dystocia



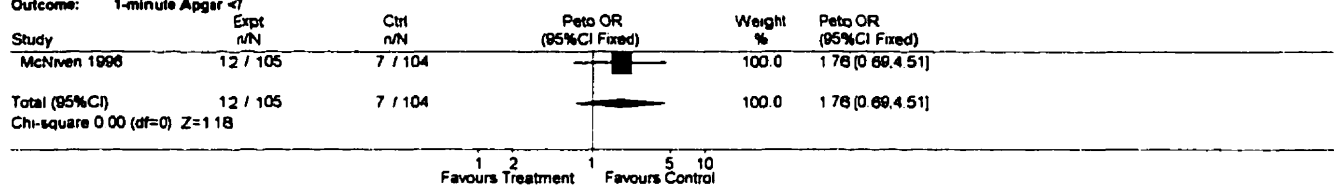
Review: Diagnosis of active labour in term pregnancy
 Comparison: Early labour assessment versus standard care
 Outcome: perceived control

43



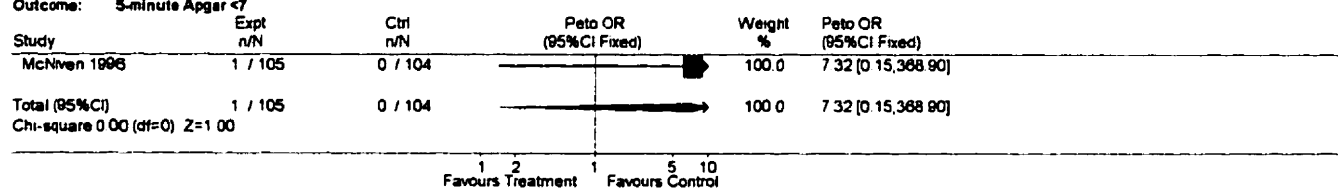
Review: Diagnosis of active labour in term pregnancy
 Comparison: Early labour assessment versus standard care
 Outcome: 1-minute Apgar <7

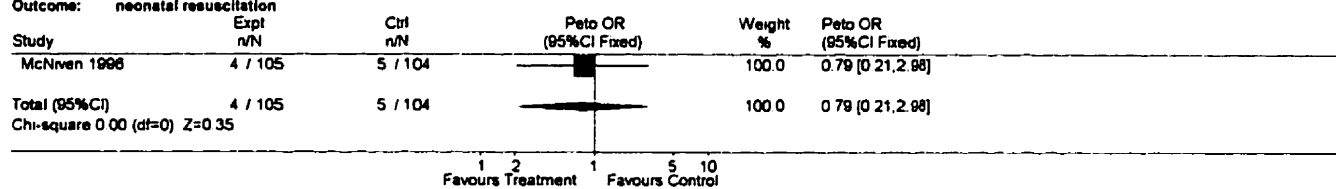
44



Review: Diagnosis of active labour in term pregnancy
 Comparison: Early labour assessment versus standard care
 Outcome: 5-minute Apgar <7

45





Antenatal Education for Self-Diagnosis of the Onset of Active Labour in Term Pregnancy

Cover Sheet

Short Title Self-diagnosis of active labour onset at term

Reviewer(s) Lauzon L, Hodnett E

Date of Most Recent Amendment 31/05/1998

Date of Most Recent Substantive Amendment 18/04/1998

Date Next Stage Expected

Contact Ms LEEANNE LAUZON RN BScN
Clinical Research Nurse
The Toronto Hospital, Western Division
EW4-556
399 Bathurst Street
Toronto
Ontario CANADA
M5T 2S8
Telephone 1: 416-603-5800 extension 2175
Telephone 2:
Facsimile: 416-603-6054
E-mail: llauzon@playfair.utoronto.ca

Sources of Support for the Review

University of Toronto

Acknowledgments

none

Conflict of Interest

none known

Antenatal Education for Self-Diagnosis of the Onset of Active Labour in Term Pregnancy

Abstract

Objective

To assess the effectiveness of teaching pregnant women specific criteria for self-diagnosis of active labour onset in term pregnancy.

Search Strategy

The register of clinical trials maintained and updated by the Cochrane Pregnancy and Childbirth Group and the Cochrane Controlled Trials Register.

Selection Criteria

All randomised controlled trials have been considered, which compare a structured antenatal education intervention for the identification of symptoms for self-diagnosis of active labour, with usual care in whatever way usual care is defined in the setting.

Data Collection and Analysis

Performed by the authors from published articles.

Main Results

The single included trial demonstrated that a specific antenatal education program was effective in reducing the mean number of visits to the labour suite before the onset of active labour. It is unclear, however, whether this intervention resulted in fewer women being sent home because they were not in labour.

Conclusions

It is unclear from the available evidence whether women would benefit from learning a specific set of criteria for self-diagnosis of active labour.

Background

Timely diagnosis of progressive labour is problematic for caregivers and expectant women. The erroneous diagnosis of active labour may lead to a subsequent diagnosis of labour dystocia, the treatments for which are associated with risks for a mother and her infant (Thornton and Lilford, 1994; Fraser, Krauss, Brisson-Carrol, Thornton, and Breart, 1995). Mothers' confidence in their caregivers may be undermined and perceptions of the birth experience negatively affected when an incorrect labour diagnosis is amended (Simkin, 1996). There may be additional financial burdens placed on facilities who assess women's labour status over multiple visits. These potential costs to women, their infants, and to health care may be avoided if admission to hospital for labour care occurs when active labour is established (Crowther et al., 1989).

Antenatal education of women and their families may be a means of imparting information to recognise active labour. This may, in turn, reduce the number of erroneous labour diagnoses by enabling women to remain out of hospital until active labour is likely to have become established. A multitude of educational resources have been developed for pregnant women. Commonly it is physicians, midwives or antenatal educators who provide information on the recognition of labour onset.

A specific program designed to teach women to recognise active labour may be beneficial to them, through potentially decreasing the incidence of early admission to hospital,

increasing women's confidence, and decreasing their anxiety. The aim of this review is to determine the effects of teaching pregnant women a specific set of criteria for diagnosing the onset of active labour. A related review focuses on the effectiveness of the application of strict criteria for labour diagnosis by caregivers [Caregivers' Use of Strict Criteria for the Diagnosis of Active Labour in Term Pregnancy].

Objectives

To assess the effectiveness of teaching pregnant women specific criteria for self-diagnosis of active labour onset in term pregnancy. The main outcomes of interest are:

- Caesarean section rate,
- admissions to labour wards or visits to labour assessment units,
- use of oxytocics, analgesics, and other intrapartum interventions,
- mothers' evaluations of their birth experiences,
- rates of hospital discharge diagnoses of "not in labour" or "false labour"
- rates of out-of-hospital emergencies (e.g. unplanned out-of-hospital births),
- admission rates to special care baby unit/neonatal intensive care unit.

Materials and Methods

Criteria for Considering Studies for This Review

Types of studies.

Randomised controlled trials which compared a structured antenatal education intervention for the identification of symptoms for self-diagnosis of active labour, with standard care in whatever way standard care was defined in the setting; violations of allocated management not sufficient to materially affect outcomes; missing data insufficient to materially affect the comparison.

Types of participants.

All pregnant women.

Types of interventions.

Any antenatal education programs specifically aimed at the identification of symptoms leading to self-diagnosis of active labour.

Types of outcome measures.

The main outcomes of interest were: rates of Caesarean section; rates of other intrapartum interventions; intrapartum complications; labour augmentation rates; degree of patient satisfaction; rates of hospital discharge diagnoses such as "not in labour" and "false labour"; rates of out-of-hospital emergencies; and neonatal outcomes, including admission rates to neonatal intensive care.

Search Strategy for Identification of Studies

This review has drawn on the search strategy developed for the Pregnancy and Childbirth Group as a whole.

Relevant trials were identified in the Group's Specialised Register of Controlled Trials and in the Cochrane Controlled Trials Register. A manual search of the Group's identified resources was conducted for the year immediately preceding the date of the most recent substantive amendment in the attempt to capture all relevant materials. See Review Group's details for more information.

Combinations of search terms used in the manual search and that of the Cochrane Controlled Trials Register for this review included the following:

- labour and education

- labour and recognise
- labour and onset
- labour and active
- triage
- labour and diagnosis
- labour and false
- labour and caregivers
- labour and nurse
- birth and education
- labour or pregnancy
- antenatal classes
- prenatal classes
- admission instructions
- hospital admission
- admission criteria
- physician and education

Methods of the Review

The reviewers independently selected and assessed the single trial resulting from the search. Names of authors, related institutions, journals of publication, and study results were known by the reviewers when inclusion criteria were applied.

Trials under consideration were evaluated for methodological quality and appropriateness for inclusion, regardless of results and conclusions, using standard Cochrane criteria. No identified trials were excluded from this review. Included trial data were processed as described in: Mulrow CD, Oxman AD (eds.), Cochrane Collaboration Handbook [updated 1 March 1997]. In: The Cochrane Library [database on disk and CD-ROM]. The Cochrane Collaboration. Oxford: Update Software; 1996-. Updated quarterly.

Description of Studies

See Characteristics of Included Studies.

Only one trial met the inclusion criteria. The trial compared a structured antenatal education intervention for the identification of symptoms for self-diagnosis of active labour, with no specific education, in an urban community hospital in the United States. Study participants were predominantly low-income single African-American women.

Methodological Quality of Included Studies

In the single trial included in this review, the method of randomisation is unclear and 15% of the sample was lost to follow-up.

Results

The study by Bonovich (1990) demonstrated that a specific antenatal education program was effective in reducing the mean number of visits to the labour suite before the onset of active labour (WMD = -0.290, 95% CI -0.0469, -0.111).

Discussion

The method of randomisation is unclear in the single trial included in this review, and so results must be considered with some caution. Attempts to contact the principal investigator for the purposes of clarification thus far have been unsuccessful.

This type of outcome measurement reporting is of limited clinical value. There is no conclusive evidence of benefit for teaching women a specific antenatal education program for self-diagnosis of active labour at present. Additionally, there is limited generalisability of results as the women participating were primarily single, low-income, urban African-Americans, in one hospital-based clinic in the US.

Conclusions

Implications for Practice

No implications for practice are warranted in light of the small amount of available evidence.

Implications for Research

Most women receive some instruction or advice regarding the signs and symptoms of labour. Whether women would benefit from learning a specific set of criteria for self-diagnosis of active labour remains unclear. It is questionable whether the potential risks and benefits of a structured educational program are of sufficient importance to warrant a large clinical trial.

Characteristics of Included Studies

Antenatal education for self-diagnosis of the onset of active labour in term pregnancy

Study	Method	Participants	Interventions	Outcomes	Notes
Bonovich 1990	Method of randomisation is unclear: 'Study numbers were sequentially assigned and used to randomise subjects to the experimental or control group.' Indicates potential for selection bias. There is potential for attrition bias as 15% of women were lost to follow-up, however the study and control groups were comparable for the analysis. Detection bias may be present as the investigator performed data collection and analysis, but as there was a single straightforward outcome, this is probably negligible. There was no performance bias.	245 nulliparous women > 16 years of age at 30+ weeks' gestation who were able to communicate effectively in English, at a single U.S. hospital-based outpatient obstetrical clinic. 37 (15%) were lost to follow-up, due to incomplete hospital records or admission to hospital for complications before the onset of normal full-term labour.	When participants had reached 37 weeks' gestation, interviews were conducted with the investigator to determine knowledge gained from family and friends regarding labour onset. Correct information was positively reinforced. Specific teaching re: palpation of uterine fundus, differentiation between Braxton-Hicks and active labour contractions, timing of contractions, recognition of amniotic fluid, and pain perception. Teaching was reinforced at subsequent weekly antenatal visits.	mean number of visits to labour suite before onset of active labour (ie discharged undelivered).	

References to Studies Included in This Review

Bonovich L. Recognizing the onset of labour. JOGNN. 1990; 19:2:141-145.

Additional References

Crowther C, Enkin MW, Keirse MJNC, and Brown I (1989). Monitoring the progress of labour. In Chalmers I, Enkin MW, and Keirse MJNC (eds.), *Effective Care in Pregnancy and Childbirth* (v.2). Oxford University Press.

Fraser WD, Krauss I, Brisson-Carrol G, Thornton J, Breart G. Amniotomy to shorten spontaneous labour. In: Neilson JP, Crowther CA, Hodnett ED, Hofmeyr GJ, Keirse MJNC (eds.) *Pregnancy and Childbirth Module of The Cochrane Database of Systematic Reviews*, [updated 03 June 1997]. Available in The Cochrane Library [database on disk and CDROM]. The Cochrane Collaboration; Issue 3. Oxford: Update Software; 1997. Updated quarterly.

Simkin P (1996). The experience of maternity in a woman's life. *JOGNN*; 25(3); 247-252.

Thornton JG and Lilford RJ (1994). Active management of labour: Current knowledge and research issues. *British Medical Journal*; 309: 366-9.

Table of Comparisons

01.00.00	Labour diagnosis education versus standard care
01.01.00	visits to labour suite before active labour
01.02.00	out-of-hospital emergencies
01.03.00	artificial rupture of membranes
01.04.00	intrapartum analgesia
01.05.00	intrapartum oxytocics
01.06.00	forceps/vacuum extraction
01.07.00	Caesarean section rates (overall)
01.08.00	Caesarean section rates for labour dystocia
01.09.00	satisfaction with care
01.10.00	admission to neonatal intensive care

Review: Self-diagnosis of active labour onset at term**Comparison or Outcome****Peto Odds Ratio****WMD (95%CI)****Labour diagnosis education versus standard care**

(95%CI)

visits to labour suite before active labour

out-of-hospital emergencies

artificial rupture of membranes

intrapartum analgesia

intrapartum oxytocics

forceps/vacuum extraction

Caesarean section rates (overall)

Caesarean section rates for labour dystocia

satisfaction with care

admission to neonatal intensive care

Not estimable

Not estimable

Not estimable

Not estimable

Not estimable

Not estimable

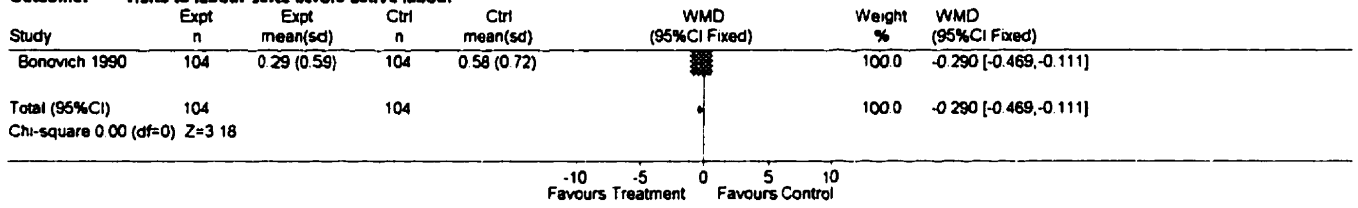
Not estimable

Not estimable

Not estimable

.1 .2 1 5 10 -10 -5 0 5 10

Review: Self-diagnosis of active labour onset at term
Comparison: Labour diagnosis education versus standard care
Outcome: visits to labour suite before active labour



CHAPTER 3

Conclusions

Through the pooling of results, a systematic review serves to increase the power of smaller studies and of those in which a treatment effect was not statistically significant (Mulrow, 1994). Unfortunately each of the systematic reviews in this theses contained only one study. Even when a study is of excellent methodological quality, as is the work of McNiven (1996), caution must be exercised in making inferences based on a single relatively small trial. A well-designed small study exhibiting favourable results may serve as a starting point for further research. Certainly when a trial's method and outcome measurements are of questionable value, as in the case of Bonovich's (1990) work, one cannot draw meaningful conclusions. Nonetheless, some inferences may be made from the available data, and implications for practice and research suggested.

McNiven (1996) demonstrated in her study of caregivers' use of specific diagnostic criteria for active labour that women in the experimental group were less likely to receive intrapartum oxytocics, and were less likely to undergo Caesarean section for labour dystocia. Women also reported higher levels of control when they were assessed prior to admission to the labour and delivery unit. The design served to eliminate or minimise potential sources of bias, and outcome measures were clinically relevant regardless of the fact that many did not achieve statistical significance, probably because of insufficient power. Despite the relatively small study size ($n=209$), this study serves to illustrate potential benefits of having caregivers employ strict diagnostic criteria for the onset of active labour. A large multi-centre

randomised controlled trial would serve to provide more conclusive evidence of the potential benefits and risks of such an intervention.

McNiven (1996) offers two alternative theoretical explanations for the results of her study, apart from the research intervention. These include the iatrogenic effects of hospitalisation in early labour and the use of epidural analgesia. It has been demonstrated that the birth environment affects psychological and physiological variables (Lederman, Lederman, Work, & McCann, 1981), and it is thought that the latent phase of labour in particular is sensitive to environmental factors (Angelini, 1986). Hodnett (1989), investigated the effects of their birth environments on childbirth outcomes for a group of 160 women. Those who gave birth at home reported significantly higher levels of control than did those who had hospital births. Although Hodnett did not focus specifically on the latent phase of labour, her findings support the notion that the hospital environment can negatively affect women's perception of control, anxiety and physiological outcomes of labour such as duration and fetal heart rate patterns (Hodnett, 1989; Lederman *et al.*, 1985). The experience of pain combined with anxiety can stimulate hormonal, vascular and other physiological symptoms which may account for unexplained abnormal labour patterns (Lederman *et al.*). Admission to hospital in early labour, as occurred with the women in the control group, may have increased maternal anxiety and distress-related thoughts (McNiven). This type of response may explain the differences in outcome measurements for each group.

McNiven (1996) also cites the more prevalent use of epidural analgesia by women in the control group as a potential factor in explaining the observed differences in Caesarean

section rates, duration of labour and use of oxytocin for labour augmentation. There would appear to be an effect of analgesia on latent labour that prolongs this phase (Berg & Rayburn, 1992), possibly through the relaxation of pelvic floor muscles which subsequently affects the dynamics of labour (Howell, 1994). Howell conducted a meta-analysis of 10 randomised controlled trials of varying methodological quality to compare epidural versus non-epidural analgesia in labour. It was found that epidural use was associated with increased duration of the first stage of labour, increased need for oxytocin, a doubling of the incidence of fetal malposition, a four-fold increase in the use of instrumental delivery as well as the increased need for Caesarean section for failure to progress. While epidural use in early labour may help to explain group differences in McNiven's trial, it is difficult to make causal inferences based on the use of epidural analgesia in labour due to the many factors involved in the use of any of the interventions.

Bonovich's (1990) study examined the effects of teaching a specific set of diagnostic criteria for the onset of active labour to pregnant women on whether or not women were sent home undelivered. The report of this study does not clearly state the method of group assignment. The single outcome measurement is described as a decreased mean number of visits to the labour suite in the experimental group before the onset of labour, or discharges undelivered. This method of reporting makes it difficult to conclude whether the intervention produced the desired effect, as it is unclear whether fewer women were discharged undelivered. These results suggest that this teaching strategy was effective; however it is not possible to conclude this based on the available evidence. Considering that the majority of women receive some form of instruction on labour recognition from care providers and other

sources of such information, it may be futile to attempt to isolate the effects of a specific curriculum when the results may not be clinically important. It is arguable whether this type of intervention warrants further investigation.

A principal objective of Bonovich's (1991) study was to see whether a low-cost nursing intervention could decrease overall the time required by labour and delivery unit staff to complete a routine procedure, through discouraging repeat labour assessment visits. She states that the rationale for this investigation was the considerable hospital operational costs for repeat assessments, in addition to having the attention of caregivers diverted from those women in active labour, giving birth, and recovering from the birth experience. Despite this motivation, the effect on operational costs was not an outcome measurement. It is clear that no consideration was given to the potential needs for support and encouragement for women not yet in active labour, and it was intimated that actively labouring women are more in need of and deserving of care.

Many hospital-based obstetrical units now have labour assessment areas wherein women are seen prior to admission for active labour care. This has occurred despite the lack of clear evidence to support their use, and that the benefits and risks associated with such areas are uncertain. It may be assumed that many of these units are not designed to provide women with any considerable amount of supportive care if they are deemed to not be in active labour. The issue, then, becomes not only women should be kept out of hospital to avoid potentially unnecessary and risky interventions, but also what is it that motivates them to present for labour assessment apart from uterine contractions. Simply meeting specific labour

criteria would appear to be insufficient to determine the care needs of a potentially labouring woman.

Both the work of McNiven(1996) and of Bonovich(1990) explore potential means by which admission for labour care may be delayed until true progressive labour is established. McNiven provided some support, encouragement, and advice for the women in the study group who did not qualify for admission and active labour care. Admission of a woman to an obstetrical ward for active labour care carries with it an implicit understanding that her labour will progress, and that non-progression, despite whether labour is truly established, is typically treated with any variety of interventions. The needs of women who have not yet met a set of criteria by which active labour is diagnosed are not well-documented and would appear to have been largely ignored.

Teaching women and reviewing with their caregivers specific criteria by which active labour is diagnosed are complementary interventions and would appear to offer some benefit to women, through enabling them to avoid potentially unnecessary and risky interventions. It is not possible to conclude that there are potential harmful effects of applying such criteria, based on the available evidence. These types of outcomes have not been considered and evaluated. It is suggested that future research efforts should include a means of describing what are the needs of women in early labour, and how best to meet them. Given the apparent proliferation of labour assessment units, the findings of this type of research could lend credence to their existence and benefit women who are deemed to not be in labour.

A benefit to conducting a systematic review is that the method, when strictly applied, enables the reviewer to envision a complete picture of the chosen research area, and to fully

comprehend the depth and breadth of research that has been conducted. Through this systematic process the true findings and the gaps in the available literature become evident, and the determination of outstanding practice and research issues is facilitated. A systematic review in this instance may be used as a catalyst for further research.

A limitation of conducting a systematic review lies with the research literature itself. Even the most well-formulated question cannot be answered without good scientific data. Evidence may be of such poor or questionable methodological quality so as to render it useless to the process of answering the review question. If the research is non-existent, the reviewer is left with the original question, which may itself become the basis for scientific investigation. In this instance, a systematic review may serve as a blueprint upon which further research may be based.

References

Arulkumaran, S., Michelsen, J., Ingemarsson, I., and Ratnam, S. S. (1987). Obstetric outcome of patients with a previous episode of spurious labour. American Journal of Obstetrics and Gynecology, 157(1): 17-20.

Baskett, T. F. (1991). Essential Management of Obstetrical Emergencies (2nd ed). Clinical Press Ltd.: Bristol, UK.

Canadian Consensus Panel on Cesarean Birth: Consensus Conference Report (1986). Indications for cesarean section: final statement of the panel of the national consensus conference on aspects of cesarean birth. Canadian Medical Association Journal, 134: 1348-1352.

Cardozo, L. and Studd, J. (1985). Abnormal labour patterns. In J. Studd (Ed.) The Management of Labour. Blackwell Scientific Publications: London.

Crowther, C., Enkin, M. W., Keirse, M. J. N. C., & Brown, I. (1989). Monitoring the progress of labour. In I. Chalmers, M. W. Enkin & M. J. N. C. Keirse (Eds.), Effective Care in Pregnancy and Childbirth (v.2). Oxford University Press: Oxford.

Fraser, W. D., Krauss, I., Brisson-Carrol, G., Thornton, J., & Breart, G. (1995). Amniotomy to shorten spontaneous labour. In: Neilson, J.P., Crowther, C.A., Hodnett, E.D., Hofmeyr, G.J., and Keirse, M.J.N.C. (eds.) Pregnancy and Childbirth Module of The Cochrane Database of Systematic Reviews, [updated 03 June 1997]. Available in The Cochrane Library [database on disk and CDROM]. The Cochrane Collaboration; Issue 3. Oxford: Update Software; 1997. Updated quarterly.

Fraser, W. D., Marcoux, S., Moutquin, J. M., Christen, A., & The Canadian Early Amniotomy Study Group (1993). Effect of early amniotomy on the risk of dystocia in nulliparous women. The New England Journal of Medicine, 328(16): 1145-1149.

Friedman, E. (1967). Labour: Clinical Evaluation and Management, New York: Appleton-Century-Crofts.

Friedman, E. (1978). Labour: Clinical Evaluation and Management, Second Edition, New York: Appleton-Century-Crofts.

Friedman, E. (1989). Normal and dysfunctional labour. In Cohen, W. R., Acker, D. B., and Friedman, E. A. (eds.) Management of Labour (2nd ed.). Aspen Publishers: Rockville, MD.

Hartley, B. M. and O'Connor, M. E. (1996). Evaluation of the 'Best Start' breast-feeding education program. Archives of Pediatric and Adolescent Medicine, 150: 868-871.

Hemminki, E. & Simukka, R. (1986). The timing of hospital admission and progress of labour. European Journal of Obstetrics, Gynaecology & Reproductive Biology, 22: 85-94.

Hodnett, E. (1989). Personal control and the birth environment: Comparisons between home and hospital settings. Journal of Environmental Psychology, 9: 207-216.

Kramer, M. S. (1996). Nutritional advice in pregnancy. In: J. P. Neilson, C. A. Crowther, E. D. Hodnett, G. J. Hofmeyr, & M.J.N.C. Keirse (Eds.), Pregnancy and Childbirth Module of The Cochrane Database of Systematic Reviews, [updated 03 June 1997]. Available in The Cochrane Library [database on disk and CDROM]. The Cochrane Collaboration; Issue 3. Oxford: Update Software; 1997. Updated quarterly.

Lederman, R., Lederman, E., Work, B. & McCann, P. (1981). Maternal psychosocial and physiologic correlates of fetal-newborn health status. American Journal of Obstetrics and Gynaecology, 139: 956-958.

Lederman, R., Lederman, E., Work, B., & McCann, P. (1985). Anxiety and epinephrine in multiparous women in labour: Relation to duration of labour and fetal heart rate pattern. American Journal of Obstetrics and Gynaecology, 153: 820-827.

McNiven, P. (1996). Feasibility study for a multi-centre randomized trial of the effect of an early labour assessment program on the rate of Cesarean delivery. Unpublished doctoral dissertation, University of Toronto, Toronto, Ontario, Canada.

McNiven, P. S., Williams, J. I., Hodnett, E., Kaufman, K., and Hannah, M. E. (1998). An early labour assessment program: A randomised, controlled trial. Birth, 25(1); 5-10.

Mulrow, C. D. (1994). Rationale for systematic reviews. British Medical Journal, 309: 597-9.

Mulrow, C. D. and Oxman, A. D. (Eds). Cochrane Collaboration Handbook [updated 1 March 1997]. In: The Cochrane Library [database on disk and CDROM]. The Cochrane Collaboration. Oxford: Update Software; 1996-. Updated quarterly.

O'Driscoll, K., Foley, M. & MacDonald, D. (1986). Active management of labour as an alternative to cesarean section for dystocia. Obstetric Gynaecology, 157: 448-451.

O'Driscoll, K., Meagher, D., & Boylan, P. (1993). Active Management of Labour: The Dublin Experience (3rd ed.).

Oxorn, H. (1986). Oxorn-Foote Human Labour and Birth. Appleton-Century-Crofts; East Norwalk, CT.

- Peisner, D. B., and Rosen, M. G. (1986). Transition from latent to active labour. Obstetrics and Gynecology, 68(4): 448-451.
- Porreco, R. P. and Thorp, J. A. (1996). The Cesarean birth epidemic: Trends, causes, and solutions. American Journal of Obstetrics and Gynecology, 175: 369-74.
- Quinn, M., Murphy, A., and Gallagher, J. (1984). Spurious labour: Does it matter? Australian/New Zealand Journal of Obstetrics and Gynaecology, 21: 1967.
- Rosen, M. G. (1990). Management of Labour: Physician Judgment and Patient Care. Science Publishing Co., Inc.: Elsevir, NY.
- Simkin, P. (1989). The Birth Partner, Boston: The Harvard Common Press.
- Simkin, P. and Enkin, M. (1989). Antenatal classes. In I. Chalmers, M. W. Enkin & M. J. N. C. Keirse (Eds.), Effective Care in Pregnancy and Childbirth (v.1). Oxford University Press: Oxford.
- Studd, J. (1973). Partograms and nomograms of cervical dilatation in management of primigravid labour. British Medical Journal, 4: 451-455.
- Simkin, P. (1996). The experience of maternity in a woman's life. Journal of Obstetrical, Gynecological and Neonatal Nursing, 25(3): 247-252.
- Stewart, P., Dulberg, C., Arnill, A., Elmslie, T., & Hall, P. (1990). Diagnosis of dystocia and management with cesarean section among primiparous women in Ottawa-Carleton. Canadian Medical Association Journal, 142(5): 459-463.
- Tay, S. K. (1991). Spurious labour: A high risk factor for dysfunctional labour and fetal distress. International Journal Gynaecology Obstetrics, 36: 209-213.

Thornton JG and Lilford RJ (1994). Active management of labour: Current knowledge and research issues. British Medical Journal, 309: 366-9.

APPENDICES

Appendix A:

General Review Method and Search Strategy for the Cochrane Pregnancy and Childbirth Group's Specialised Registry of Controlled Trials

1. INTRODUCTION

Healthcare providers, consumers, researchers, and policy makers are inundated with unmanageable amounts of information. We need systematic reviews to efficiently integrate valid information and provide a basis for rational decision making {188}. Systematic reviews establish where the effects of healthcare are consistent and research results can be applied across populations, settings, and differences in treatment (e.g. dose); and where effects may vary significantly. The use of explicit, systematic methods in reviews limits bias (systematic errors) and reduces random errors (simple mistakes), thus providing more reliable results upon which to draw conclusions and make decisions {103, 233}. Meta-analysis, the use of statistical methods to summarise the results of independent studies, can provide more precise estimates of the effects of healthcare than those derived from the individual studies included in a review {335, 177, 70, 439}.

Wider recognition of the key role of reviews in synthesising and disseminating the results of research has prompted people to consider the validity of reviews. In the 1970s and early 1980s, psychologists and social scientists drew attention to the systematic steps needed to minimise bias and random errors in reviews of research {43, 24, 440, 37, 441}. It was not until the late 1980s that people drew attention to the poor scientific quality of healthcare review articles {189, 197, 361}. However, recognition of the need for systematic reviews of healthcare has grown rapidly and continues to grow, as reflected by the number of articles about review methods {442}, the number of systematic reviews published in healthcare journals {443}, and the rapid growth of the Cochrane Collaboration {444}.

Recognition of the importance of systematic reviews has also stimulated a growing number of empirical studies of the methods used in reviews {445}.

This Handbook builds on the work of a large number of people, including those represented in the Cochrane Review Methodology Database {442}, input from Cochrane Methods Working Groups {444}, practical experience and feedback from Collaborative Review Groups {444} which have taken on the daunting task of systematically reviewing the effects of healthcare within their areas of interest, and Cochrane Centres {444} which provide training for reviewers. Whenever possible recommendations made here are based on empirical evidence and advice from Cochrane Methods Working Groups.

Our aim is to help reviewers make good decisions about the methods they use relative to the specific healthcare problems that they address, rather than dictate arbitrary standards. The guidelines provided here are intended to help reviewers to be systematic and explicit (not mechanistic!) about the questions they pose and how they derive answers to those questions. These guidelines are not a substitute for good judgment.

The Cochrane Collaboration and the Handbook focus particularly on systematic reviews of randomised controlled trials (RCTs) because they are likely to provide more reliable information than other sources of evidence on the differential effects of alternative forms of healthcare. Systematic reviews of other types of evidence can also help those wanting to make better decisions about healthcare, particularly forms of care where RCTs have not been done and may not be possible or appropriate. The basic principles of reviewing research are the same, whatever type of evidence is being reviewed. Although we focus mainly on systematic reviews of RCTs we address issues specific to reviewing other types of evidence when this is relevant.

Cochrane Reviews have a standard format that we describe in the next section (section 2). Those preparing a review should begin by developing a protocol (Section 3). The seven succeeding sections are organised according to the steps of preparing and maintaining a systematic review:

- Formulating the problem
- Locating and selecting studies
- Critical appraisal of studies
- Collecting data
- Analysing and presenting results
- Interpreting results
- Improving and updating reviews

In the last section we take up specific issues about using individual patient data in reviews.

10.6 Using rigorous review methods

It is neither feasible nor desirable to dictate the decisions that a reviewer should take. These will vary from review to review depending on the topic, the nature of the available evidence and the resources available to the reviewer. However, in general, the validity of Cochrane Reviews is ensured by:

- Searching as thoroughly as possible for studies meeting the inclusion criteria of a review, relying as much as possible on centralised efforts to assist with this and ensure the thoroughness and efficiency with which RCTs are identified
- Use of explicit criteria for selecting trials for inclusion in a review and for assessing the quality of included trials
- Application of these criteria by more than one reviewer where appropriate and feasible, to ensure the reproducibility of the judgments that are made
- Ongoing efforts to collect missing information that might contribute importantly to a review, to the extent possible depending on the availability of resources and data
- Collection of individual patient data from trialists where appropriate and feasible, to the extent possible depending on the availability of resources and data
- Use of appropriate statistical techniques, where appropriate, to synthesize results
- Use of sensitivity analyses to test the robustness of the results of a review relative to any judgments or assumptions
- Cautious use of subgroup analyses and avoidance of over-interpretation of any subgroup analyses that are undertaken
- Carefully drawn conclusions, including implications for practice and future research, based on cautious interpretation of results - taking into account the limitations of the review and variability in the values and conditions of those making decisions
- Full reporting of the materials and methods used in undertaking the review

Just as it is possible to update Cochrane Reviews in the light of new evidence, it is possible to improve upon the methods. Moreover, because the methods are explicitly reported in Cochrane Reviews, users can judge for themselves the validity of the results of a review.

Search strategy for specialised register

THE COCHRANE PREGNANCY AND CHILDBIRTH GROUP'S SPECIALISED REGISTER OF CONTROLLED TRIALS

1. INTRODUCTION

The editorial team of the Cochrane Pregnancy and Childbirth Group assembles, maintains and administers centrally a specialised register of controlled trials as a service to the 86 reviewers who are members of the Group. The basis for the register was established between 1978 and 1985 with the creation of the Oxford Database of Perinatal Trials (ODPT). The development and features of ODPT have been described in a series of reports published over the past 15 years (see references, below).

The specialised register created to assist members of the Cochrane Pregnancy and Childbirth Group currently contains more than 7000 records referring to completed or ongoing trials, with an annual accrual rate of about 300 new records. Reports are identified by the formal search methods described below, or by informal discovery.

On the basis of the health topic(s) and/or form(s) of care covered, every record in the register has been assigned by the editorial team to one or more reviewers in the Cochrane Group, according to the agreed spheres of responsibility of each.

2. ELIGIBILITY CRITERIA

TOPIC SCOPE: Controlled trials comparing alternative forms of care used either during pregnancy (but not to terminate early pregnancy), or within 28 days of delivery.

STUDY DESIGN: A controlled trial has been defined as a trial involving humans in which allocation to the intervention has either been at random, or by some quasi-random method, such as by alternation, or on the basis of the case record number or date of birth.

These criteria have been applied fairly liberally to avoid excluding potentially useful studies involving concurrent comparisons of alternative policies. In other words, the register includes reports which, if necessary, can subsequently be rejected as methodologically inadequate by a member of the Group preparing a systematic review.

3. SEARCH FOR ELIGIBLE CONTROLLED TRIALS

(i) Electronic search of bibliographic databases

The National Library of Medicine MEDLINE database has been searched back to 1966, and is updated quarterly. The method of access and search strategy have been adjusted from time to time. The current search strategy, using SilverPlatter MEDLINE, is as follows:

- 1 RANDOMIZED-CONTROLLED TRIAL in PT
- 2 RANDOMIZED-CONTROLLED-TRIALS
- 3 RANDOM-ALLOCATION
- 4 DOUBLE-BLIND-METHOD
- 5 SINGLE-BLIND-METHOD
- 6 CLINICAL-TRIAL in PT
- 7 explode CLINICAL-TRIALS
- 8 (clin* near trial*) in TI
- 9 (clin* near trial*) in AB
- 10 (singl* or doubl* or trebl* or tripl*) near (blind* or mask*)
- 11 (#10 in TI) or (#10 in AB)
- 12 PLACEBOS
- 13 placebo* in TI
- 14 placebo* in AB
- 15 random* in TI
- 16 random* in AB
- 17 RESEARCH-DESIGN
- 18 #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #11 or #12 or #13 or #14 or #15 or #16 or #17
- 19 explode FETUS
- 20 explode INFANT-NEWBORN
- 21 explode PREGNANCY
- 22 #19 or #20 or #21
- 23 #18 and #22
- 24 TG=ANIMAL not (TG=HUMAN and TG=ANIMAL)
- 25 #23 not #24
- 26 PT=CONTROLLED-CLINICAL-TRIAL
- 27 #18 or #26
- 28 #27 and #22
- 29 #28 not #24

This search strategy was devised by Carol Lefebvre of the UK Cochrane Centre.

(ii) Handsearch of journals

A systematic journal handsearch (for perinatal trials only) of the journals listed below has been carried out and maintained from the initial date of publication of each journal, or from 1950, whichever was the earlier. Maintenance of the search of some journals • (mainly paediatric and anaesthetic) was discontinued with the registration of other Cochrane entities during 1993 where commitments were made by others to search for trials on behalf of the Cochrane Collaboration.

Because of the history of the Pregnancy and Childbirth Group and, in particular, to maintain their published reviews, several journals continued to be search by the Group

for perinatal trials only. The maintenance of several other journals ceased at the end of 1996 ** because of the advances made by the Trials Register Development Group to facilitate the speedier notification of trial reports to Cochrane Groups. The Pregnancy and Childbirth Group is now aiming to concentrate their handsearching efforts on the searching of journals that are core to their scope on behalf of the Cochrane Collaboration as a whole. This also includes retrospective searching.

Acta Anaesthesiol Scand (and supplements): From 1st issue and continuing
 Acta Obstet Gynecol Scand (and supplements): From 1950 and continuing
 Acta Paediatr Scand*: From 1st issue through 1993
 Am J Clin Nutr: From 1st issue and continuing
 Am J Dis Child*: From 1950 through 1993
 Am J Obstet Gynecol: From 1950 and continuing
 Anaesth Intens Care: From 1st issue and continuing
 Anaesthesia: From 1950 and continuing
 Anesth Analg: From 1st issue and continuing
 Anesthesiology: From 1950 and continuing
 Arch Dis Child*: From 1950 through 1993
 Aust NZ J Obstet Gynaecol: From 1st issue and continuing
 Birth: From 1st issue and continuing
 BMJ**: From 1950 through 1996
 Br J Anaesth: From 1950 and continuing
 Br J Obstet Gynaecol: From 1st issue and continuing
 Can J Anaesth: From 1st issue and continuing
 Can Med Assoc J*: From 1950 and continuing
 Clin Pharmacol Ther: From 1st issue and continuing
 Curr Med Res Opin*: From 1st issue through 1993
 Dev Med Child Neurol*: From 1st issue through 1993
 Early Hum Dev*: From 1st issue through 1993
 Eur J Obstet Gynecol Reprod Biol: From 1st issue and continuing
 Geburtshilfe Frauenheilkd: From 1950 and continuing
 Gynecol Obstet Invest: From 1st issue and continuing
 Int J Gynaecol Obstet (and supplements): From 1st issue and continuing
 Int J Obstet Anesthesia: From October 1994 to October 1995
 JAMA**: From 1st issue through 1996
 J Am Coll Surg: From 1950 and continuing
 J Gynecol Obstet Biol Reprod (Paris): From 1st issue and continuing
 J Int Med Res*: From 1st issue through 1993
 J Nurs Midw: From 1st issue and continuing
 J Obstet Gynaecol: From 1st issue and continuing
 JOGN Nurs*: From 1st issue through 1993
 J Pediatr*: From 1950 through 1993
 J Pediatr Gastroenterol Nutr*: From 1st issue through 1993
 J Perinat Med: From 1st issue and continuing
 J Reprod Med: From 1st issue and continuing

Lancet**: From 1950 through 1996
 Med J Aust: From 1950 and continuing
 Midwifery: From 1st issue and continuing
 N Engl J Med**: From 1950 through 1996
 Nurse Res*: From 1st issue through 1993
 NZ Med J**: From 1950 through 1996
 Obstet Gynecol: From 1st issue and continuing
 Pediatr Res*: From 1st issue through 1993
 Pediatrics*: From 1950 through 1993
 Practitioner: From 1950 and continuing
 Prostaglandins*: From 1st issue through 1993
 Reg Anesth: From 1st issue and continuing
 S Afr J Obstet Gynaecol*: From 1st issue through 1993
 S Afr Med J*: From 1950 through 1993
 Surg Gynecol Obstet*: From 1950 through 1993
 Ugeskr Laeger*: From 1950 through 1993
 Z Geburtshilfe Perinatol: From 1st issue and continuing
 Zentralbl Gynakol: From 1950 and continuing

(iii) Handsearch of conference proceedings

American College of Obstetricians and Gynecologists' Annual Meeting: 36th, 37th, 39th, 40th, 41st
 Argentinian Congress of Perinatology: 3rd
 Australian Perinatal Society: 14th
 Birth Conference: 1st, 2nd, 3rd, 4th, 5th, 6th, 7th, 8th, 9th
 British Congress of Obstetrics and Gynaecology: 23rd, 25th, 26th, 27th
 British Paediatric Association Annual Meeting: 14th, 15th, 27th, 60th, 61st, 62nd, 63rd, 65th
 European Congress of Allied Specialists in Maternal and Neonatal Care: 4th
 European Congress of Obstetrical Anaesthesia and Analgesia: 1st
 European Congress of Perinatal Medicine: 5th, 6th, 8th, 10th, 11th, 12th, 14th, 15th
 European Congress on Prostaglandins in Reproduction: 1st, 2nd
 European Congress on Ultrasound in Medicine and Biology: 6th
 Federation of the Asia-Oceania Perinatal Societies' Congress: 6th, 9th
 International Confederation of Midwives Triennial Congress: 24th
 International Congress on Psychosomatic Medicine in Obstetrics and Gynaecology: 3rd, 5th
 International Society for the Study of Hypertension in Pregnancy (IISHP) European Branch: 1st, 5th, 6th
 Priorities in Perinatal Care in South Africa: 4th, 7th, 10th, 11th, 14th, 15th
 Society of Obstetricians and Gynaecologists of Canada Annual Meeting: 49th
 Society of Perinatal Obstetricians' (USA) Annual Meeting: 3rd, 6th, 7th, 8th, 9th,

10th

Society for Gynecologic Investigation (USA) Annual Program: 31st, 34th, 37th, 39th, 40th

World Congress of Perinatal Medicine: 1st, 2nd

World Congress of Gynecology and Obstetrics: 11th, 12th, 13th, 14th

World Congress of Hypertension in Pregnancy: 7th, 8th, 9th, 10th

(iv) Surveys to identify unpublished and ongoing trials

During the second half of 1986 and early 1987, letters were sent to approximately 42,000 obstetricians and pediatricians in 18 countries in an attempt to identify unpublished controlled trials in perinatal medicine. The countries included in the survey were selected because they had generated more than 90% of the published reports of controlled trials in the Oxford Database of Perinatal Trials. This resulted in the notification of 395 unpublished randomized trials. Only 18 of the trials had been completed more than 2 years before the survey, a period during which at least 2300 reports of perinatal trials had been published. Of the 395 unpublished trials, 125 had ceased recruitment within the 2 years prior to the survey, 193 were actively recruiting at the time of the survey, and 59 were about to begin recruitment.

In 1991, prompted by the disappointing response to the earlier survey of individuals in an attempt to obtain information about unpublished and ongoing trials, a further, more focussed survey was conducted of clinical and academic institutions and funding agencies in the United Kingdom and North America to assess the feasibility of voluntary registration of trials. The experience gained in this and the earlier survey suggested that publication bias could not be addressed successfully by attempts to obtain information about unpublished trials retrospectively. This has led members of the Cochrane Pregnancy and Childbirth Group to support calls for prospective registration of trials, at inception.

Appendix B:

**Critical Appraisal of Studies. From: The Cochrane Collaboration Handbook
(Version 4.0, Chapter Six)**

6. CRITICAL APPRAISAL OF STUDIES

Critical appraisal of individual studies that are summarised in systematic reviews is necessary to limit bias in conducting the systematic review, gain insight into potential comparisons, and guide interpretation of findings. Parameters that warrant appraisal are those related to applicability of findings, validity of individual studies, and certain design characteristics that affect interpretation of results. Applicability, also called external validity or generalisability by some, is related to the definition of the key components of well-formulated questions outlined in section 4. Specifically, whether a review's findings are applicable to a particular population, intervention strategy or outcome is dependent upon the studies selected for review, and on how the studies as well as the reviewers define the people, interventions and outcomes of interest.

Interpretation of results is dependent upon the validity of the included studies and other characteristics. For example, a review may summarise twenty valid trials that evaluate the effects of antiischemic agents on symptoms of chest pain in adults with prior myocardial infarction. However, the trials may examine different preparations and doses of antiischemic agents and may have varying durations. These latter issues would affect interpretation though they may not be directly relevant to the internal validity of the trials. Examples of what and how to abstract data related to applicability and design factors likely to affect the interpretation of findings will be given in the next section (section 7). The remainder of this section will focus on critically appraising the validity of individual studies included in a systematic review. As most Cochrane Reviews focus on randomised trial data, we will concentrate on how to appraise the validity of such data.

6.1 Validity

In the context of a systematic review, the validity of a study is the extent to which its design and conduct are likely to prevent systematic errors, or bias {240}. An important issue that should not be confused with validity is precision. Precision is a measure of the likelihood of random errors. It is reflected in the confidence interval around the estimate of effect from each study and the weight given to the results of each study when an overall estimate of effect or weighted average is derived using meta-analysis. More precise results are given more weight.

Variation in validity can explain variation in the results of the studies included in a systematic review. More rigorous studies may be more likely to yield results that are closer to the "truth". Quantitative analysis of results from trials of variable validity can result in "false positive" conclusions (erroneously concluding an intervention is effective) if the less rigorous studies are biased toward overestimating an intervention's effectiveness. They might also come to "false negative" conclusions (erroneously concluding no effect) if the less rigorous studies provide less precise or biased estimates of an intervention's effect {135}.

It is important to systematically complete critical appraisal of all studies in a review even if there is no variability in either validity or results of the included studies. For instance, the results may be consistent among studies but all the studies may be flawed. In this case, the systematic review's conclusions would not be nearly as strong as if a series of rigorous studies yielded consistent results about an intervention's effect.

6.2 Sources of bias in trials of healthcare interventions

There are four sources of systematic errors in trials of the effects of healthcare: selection bias, performance bias, attrition bias and detection bias (see figure below). Unfortunately, we do not have strong empirical evidence of a relationship between trial outcomes and specific criteria or sets of criteria used to assess the risk of these biases {240, 491}. There is, however, a logical basis for suspecting such relationships and good reason to assess these four potential biases {81}.

6.3 Selection bias

One of the most important biases that may distort treatment comparisons is that which can result from the way that comparison groups are assembled {486}. Using an appropriate method for preventing foreknowledge of treatment assignment is crucially important in trial design. When assessing a potential participant's eligibility for a trial, those who are recruiting participants and the participants themselves should remain unaware of the next assignment in the sequence until after the decision about eligibility has been made. Then, after assignment has been revealed, they should not be able to alter the assignment or the decision about eligibility. The ideal is for the process to be impervious to any influence by the individuals making the allocation. This will be most securely achieved if an assignment schedule generated using true randomisation is administered by someone who is not responsible for recruiting subjects, such as someone based in a central trial office or pharmacy. If such central randomisation cannot be organised, then other precautions are required to prevent manipulation of random assignment by those involved in recruitment.

The process of concealing assignment until treatment has been allocated has sometimes been referred to as "randomization blinding" {123}. This term does not clearly distinguish concealed allocation from blinding of patients, providers, outcome evaluators and analysts and is unsatisfactory for three reasons. First, the reason for concealing the assignment schedule is to eliminate selection bias. In contrast, blinding (used after allocation of treatments) reduces performance and detection biases. Second, from a practical standpoint, concealing treatment assignment up to the point of assignment is always possible, regardless of the study question, but blinding after allocation may be impossible, as in trials comparing surgical with medical treatment. Third, control of selection bias is relevant to the trial as a whole, and thus to whatever outcomes are being compared. In contrast, control of detection bias is often outcome-

specific and may be accomplished successfully for some outcomes in a trial but not others. Thus, blinding up to allocation and blinding after allocation of treatment are addressing different sources of bias, are inherently different in their practicability and may apply to different parts of a trial. To clearly distinguish these different forms and purposes of "blinding", we will refer to the process of concealing assignments as allocation concealment and reserve blinding for measures taken to reduce bias after treatment has been assigned.

Empirical research has shown that lack of adequate allocation concealment is associated with bias [123, 62]. Indeed, concealment has been found to be more important in preventing bias than other components of allocation, such as the generation of the allocation sequence (e.g., computer, random number table, alternation). Thus, trials can be judged on the reported method of allocation concealment. Information should be presented that provides some assurance that allocations were not known until the point of allocation, at least. The method for assigning participants to treatments should be robust against patient and clinician bias and its description should be clear. The following are some approaches that can be used to assure adequate concealment schemes.

- centralised (e.g., group assignment by a central office unaware of subject characteristics) or pharmacy-controlled randomisation
- pre-numbered or coded identical containers which are administered serially to participants
- on-site computer system combined with group assignments in a locked unreadable computer file that can be accessed only after entering characteristics of an enrolled subject
- sequentially numbered, sealed, opaque envelopes

Other approaches may include statements that imply an approach similar to ones listed above, along with reassurance that the person who generated the allocation scheme did not administer it. Some schemes may be innovative and not fit any of the approaches above, but still seem to provide adequate concealment.

Approaches to allocation concealment that should be considered clearly inadequate include: alternation, case record numbers, dates of birth, day of the week, and any allocation procedure that is entirely transparent before assignment, such as an open list of random numbers of assignments. When trials do not report any concealment approach, adequacy should be considered unclear. Examples include merely stating that a list or table was used, only specifying that sealed envelopes were used and reporting an apparently adequate concealment scheme in combination with other information that leads the reviewer to be suspicious. When reviewers enter studies into Review Manager (RevMan) they are required to whether allocation concealment was

adequate (A), unclear (B), inadequate (C), or that allocation concealment was not used (D) as a criterion to assess validity.

6.4 Performance bias

Performance bias refers to systematic differences in care provided to comparison groups other than the intervention of interest. To protect against unintended differences in care and placebo effects, those providing and receiving care can be "blinded" so that they do not know the group to which the recipients of care have been allocated. Some research suggests that such blinding is indeed important in protecting against bias {62, 487, 488}. Studies have shown that contamination (provision of the intervention to the control group) and cointervention (provision of unintended additional care to either comparison group) can affect study results {489, 490}. Furthermore, there is evidence that participants who are aware of their assignment status report more symptoms, leading to biased results {488}. For these reasons, reviewers may want to consider the use of "blinding" as a criterion for validity. This can be done with the following questions: Were the recipients of care unaware of their assigned treatment? Were those providing care unaware of the assigned therapy?

A third question addressing blinding and detection bias is often added: Were persons responsible for outcome assessments unaware of the assigned therapy? This addresses detection bias, as noted below.

Blinding is likely to be particularly important in research with subjective outcome measures such as pain {62, 487, 488}. Reviewers working on topics where blinding is likely to be important may want to develop specific criteria for judging the appropriateness of the method that was used for binding. In some areas it may be desirable to use the same criterion across reviews, in which case a review group might want to agree to a standard approach for assessing blinding {89, 62, 480, 491}.

6.5 Attrition bias

Attrition bias refers to systematic differences between groups in losses of participants from the study. It has sometimes been referred to as exclusion bias but we call it attrition bias to prevent confusion with pre-allocation exclusion and inclusion criteria for enrolling people. Because of inadequacies in reporting how losses of participants (e.g., withdrawals, dropouts, protocol deviations) are handled, reviewers should be cautious about implicit accounts of follow-up. The approach to handling losses has great potential for biasing the results and reporting inadequacies cloud this problem. What is reported, or more frequently implied, in trial reports on attrition after allocation has not been found to be consistently related to bias {62}. Thus reviewers should be cautious about using reported follow-up as a validity criterion, particularly when it is implied rather than explicitly reported. This is a general recommendation, however,

and may not apply to certain topic areas that have higher quality reporting or where it is possible to obtain missing information from authors.

6.6 Detection bias

Detection bias refers to systematic differences in outcome assessment. Trials that blind outcome assessors regarding treatment allocation should logically be less likely to be biased than trials that do not. However, at least two studies have failed to demonstrate empirically a relationship between blinding of outcome assessment and study results, possibly due to inadequacies in trial reports {62, 245}.

Somewhat different from bias in outcome assessment is bias due to selective reporting of results. This source of bias may be important in areas where multiple outcome measures are used, as in evaluations of treatments for rheumatoid arthritis {157}. Therefore, reviewers may want to consider specification of predefined primary outcomes and analyses by the investigators indicators of validity. Alternatively, selective reporting of results could be taken to suggest the need for better reporting and efforts by reviewers to obtain missing data.

6.7.1 Simple approaches

There are several ways to rate validity. One is to rate individual criteria as "met", "unmet", or "unclear" and to use individual criteria, such as adequacy of allocation concealment, in sensitivity analyses (see section 8.8). However, having used several explicit criteria to assess validity, it is desirable to summarise these somehow to derive an overall assessment of how valid the results of each study are. A simple approach to doing this is to use three categories such as the following:

	Interpretation	Relationship to individual criteria
A Low risk of bias	Plausible bias unlikely to seriously alter the results	All of the criteria met
B Moderate risk of bias	Plausible bias that raises some doubt about the results	One or more criteria partly met
C High risk of bias	Plausible bias that seriously weakens confidence in the results	One or more criteria not met

The relationships suggested above will most likely be appropriate if only a few assessment criteria are used and if all the criteria address only substantive, important

threats to the validity of results. In general and when possible, reviewers should obtain further information from the authors of a report when it is unclear whether a criterion was met.

6.7.2 "Quality" scales and checklists

David Moher and his colleagues have identified 25 scales and 9 checklists that have been used to assess the validity and "quality" of randomised controlled trials {240, 491}. These scales and checklists include anywhere from 3 to 57 items and take from 10 to 45 minutes to complete. Almost all of the items in the instruments are based on suggested or "generally accepted" criteria that are mentioned in clinical trial textbooks. Many of the instruments are liable to confuse the quality of reporting with the validity of the design and conduct of a trial. Moreover, scoring is based on whether something was reported (such as how participants were allocated) rather than whether it was done appropriately. Many also contain items that are not directly related to validity, such as whether a power calculation was done (an item that relates more to the precision of the results) or whether the inclusion and exclusion criteria were clearly described (an item that relates more to applicability than validity).

Because there is no "gold standard" for the "true" validity of a trial, the possibility of validating any proposed scoring system is limited. While it is possible to apply basic principles of measurement to the development of a scale for assessing the validity of randomised controlled trials, the relationship between such a score and the degree to which a trial is free from bias is not obvious. None of the currently available scales for measuring the validity or "quality" of trials can be recommended without reservation. If reviewers or review groups choose to use such a scale, it must be with caution.

Most of the available scales for assessing the validity of randomised controlled trials derive a summary score by adding the scores (with or without weights) for each item. While this approach offers appealing simplicity, it is not supported by empirical evidence {62, 146}. Notably, scales with multiple items and complex scoring systems take more time to complete than simple approaches. They have not been shown to provide more reliable assessments of validity. They may carry a greater risk of confusing the quality of reporting with the validity of the trial. They are more likely to include criteria that do not directly measure internal validity, and they are less likely to be transparent to users of the review. For these reasons, it is preferable to use simple approaches for assessing validity that can be fully reported (i.e. how each trial scored on each criterion).

6.8 Bias in non-experimental studies

The logical reason for focusing on randomised controlled trials in Cochrane Reviews is that randomisation is the only means of allocation that controls for unknown and unmeasured confounders as well as those that are known and measured. Differences

between comparison groups in prognosis, responsiveness to treatment or exposure to other factors that affect outcomes can distort the apparent magnitude of effects of the intervention of interest. It is possible to control or adjust for confounders that are known and measured in observational studies, such as case-control and cohort studies. However, it is not possible to adjust for those factors that are not known to be confounders or that were not measured. Unfortunately it can rarely, if ever, be assumed that all important factors relevant to prognosis and responsiveness to treatment are known, and for those that are known difficulties can arise in measuring and accounting for them in analyses. Empirical evidence supports these logical concerns {492}. Selection bias can distort effects in either direction, causing them to appear either larger or smaller than they are. It is generally not possible to predict the magnitude, and often not even the direction of this bias in specific studies. However, on average, selection bias tends to make treatment effects appear larger than they are and the size of these distortions can be as large or larger than the size of the effects that are being measured {492}.

Despite these concerns, there is sometimes good reason to rely on observational studies for information about the effects of healthcare interventions, and to include such studies in Cochrane Reviews. For example, well designed observational studies have provided useful data regarding the effects of interventions such as mandatory use of helmets by motorcyclists, screening for cervical cancer, dissemination of clinical practice guidelines to change professional practice and rare adverse effects of medication.

Various criteria have been suggested to critically appraise the validity of observational studies {493, 494, 495, 496}. In general, the same four sources of bias noted above can be applied to other types of comparative studies, as illustrated below:

Source of bias	Cohort studies	Case-control studies
Selection bias	Control for confounders	Matching
Performance bias	Measurement of exposure	Measurement of exposure
Attrition bias	Completeness of follow-up	Completeness of follow-up
Detection bias	Blinding	Case definition

Concerns about attrition bias are similar in trials, cohort studies and case-control studies and relate to the extent that those entered into a study are appropriately accounted for in the results. Concerns about detection bias are also similar for cohort studies, and are related to the case definition that is used in case-control studies (since people are entered into such studies based on knowledge of the outcome of interest). The major difference between trials and observational studies has to do with selection

bias and the need to identify and account for potential confounders in observational studies. To do this reviewers must make judgements about what confounders are important and the extent to which these were appropriately measured and controlled for. Assessing "performance bias" is also more difficult in observational studies since it is necessary to measure exposure to the intervention of interest and ensure that there were not differences in exposure to other factors that could affect outcomes. In addition to considerations of blinding, which are similar to those in trials, it is important to consider whether exposure was measured in a similar and unbiased way in the groups being compared. So, for example, in addition to concerns about bias due to confounders in cohort and case control studies of the effects of post-menopausal hormone replacement therapy, investigators and reviewers must ensure that use of hormones was measured in an unbiased way.

In summary, a great deal of judgement is necessary in assessing the validity of observational studies. Judgement is also needed when the validity of trials is assessed, but the nature of observational studies makes them even more difficult to critically appraise. This requires a thorough understanding of both the problem that is the focus of the review and methodological considerations. Caution is advised.

6.9 Application of critical appraisal criteria

Several basic decisions must be made regarding the critical appraisal studies, similar to those made regarding the process of selecting studies (section 5.7). A prime consideration is the number of reviewers. Should there be one or more than one? How many are necessary and how many are too many? Will reviewers review the same articles to maximise reliability or mutually exclusive sets of reports to minimise workload? A concomitant consideration is reviewers' backgrounds and whether previous training and experience in study design or critical appraisal will be required.

Conducting systematic reviews with multiple reviewers is a two-sided coin. On the one hand it may limit bias and minimise errors and improve reliability of findings, but more than one creates the potential for disagreement among reviewers. When multiple reviewers are planned, there should be an explicit procedure or decision rule identified a priori for identifying and resolving disagreement. As a general rule, we recommend that at least two reviewers assess information that involves subjective interpretation and information that is critical to the interpretation of results (e.g., outcome data). The next section (section 7) describes methods for reaching and monitoring consensus when more than one reviewer is used.

Regardless of the number of reviewers, it is important to test any assessment criteria that are planned on a pilot sample of articles to ensure that the appraisal criteria can be applied consistently. A suggested sample would be three to six papers that span a range of low to high risk bias.

Should reviewers be especially trained in research methods, the content area of a review or both? Although experts in content areas may have pre-formed opinions that can bias their assessments {233}, they may nonetheless give more consistent assessments of the validity of trials than persons without content expertise {480} and they may have valuable insights that are different than those that someone with methodological expertise alone would have. It would seem intuitively desirable to use both content experts and non-experts and to ensure that both have an adequate understanding of the relevant methodological issues.

Reviewers must also decide whether those assessing study validity will be blinded to the names of the authors, institutions, journal and results of a trial when they apply critical appraisal criteria to the methods. Some empirical evidence suggests that blind assessment of reports might produce lower and more consistent scores than open assessments {480}. However, such assessments are very time consuming. Reviewers must weigh the potential benefits of blind assessments against the costs involved in deciding whether or not to blind the reviewers. Further research is underway comparing blind and open assessments of trial validity and these results may help guide this decision.

6.10 Incorporating assessments of study validity in reviews

There are several ways in which validity assessments can be used in a review:

- as a threshold for inclusion of studies
- as a possible explanation for differences in results between trials
- in sensitivity analyses
- as weights in statistical analysis (meta-analysis) of the results

Failure to meet one or more validity criteria may indicate such a high risk of bias in some reviews that it constitutes grounds for exclusion of those studies. For example, for highly subjective outcomes such as pain, reviewers may decide to include only trials that prevent "performance bias" by blinding participants. The decision about where to set the cut point for inclusion can be conceptualised as existing on a continuum between "free from bias" and "undoubtedly biased" as illustrated below:

If reviewers raise the methodological cut-point for including studies, there will be less variation in validity among the included reports. Assessments of validity would then categorise studies by the risk of bias within the range above the cut-point for inclusion. With a sufficiently high cut-point, variation in validity among included reports may be moot.

There are several methods to examine whether validity may explain differences among study results {135}. Visual plots of the results arranged in order of their validity can be used. A second approach is to analyse subgroups of studies above a methodological cut-point, preferably one specified a priori. This approach can be used whether or not the results are heterogeneous, as a sensitivity analysis to determine if the overall results are the same when only studies with little risk of bias are included in the analysis. A third approach is to combine the results of each study sequentially in order of their assessed validity (a type of "cumulative meta-analysis"), examining the impact on the overall results as trials of decreasing validity are included (see section 8.6.2).

A fourth approach is to use statistical methods to weight studies according to their assessed validity or to use "meta-regression" to explore the relationship between validity and the magnitude of effect across studies (see section 8.6.1). Statistical methods for combining the results of studies generally weight the influence of each study by the inverse of the variance for the estimated measure of effect. In other words, studies with more precise results (narrower confidence intervals) are given more weight. It is also possible to weight studies according to validity so that more valid studies have more influence on the summary result. The main objection to this approach is that there is no empirical basis for determining how much weight to assign to different validity criteria or for quantitatively relating differences on any of the available "quality" scales to differences in the risk of bias.

It is possible using RevMan 3.0 to order studies according to either adequacy of concealment of allocation or "user defined" assessments of validity. Subgroup analyses based on assessments of validity can be done, although a test of statistical significance of differences between subgroups of studies has not been implemented. A function to facilitate sensitivity analyses also has not yet been implemented, but it is possible to do these by deleting or adding studies. RevMan does not include an option for weighting studies by methodological validity and neither cumulative meta-analysis or meta-regression is possible using RevMan 3.0.

6.11 Limitations of critical appraisal

There are two major difficulties with critically appraising the validity of studies. The first is inadequate reporting of trials {67, 213, 497}. It is possible to assume if something was not reported it was not done. However, this is not necessarily correct. Reviewers should attempt to obtain additional clarifying data from investigators, but this may be difficult. The application of standards for reporting trials {67, 497} can facilitate critical appraisal.

The second limitation, which in part is a consequence of the first, is limited empirical evidence of a relationship between parameters thought to measure validity and actual

trial outcomes. As noted above, there is empirical evidence suggesting that, on average, both inadequate concealment of allocation and lack of double blinding result in over-estimates of the effects of treatment. Clearly much more research needs to be done to establish which criteria for assessing validity are indeed important determinants of study results and when. Improved reporting of methods will also facilitate such research. Meanwhile, reviewers should avoid the use of "quality scores" and undue reliance on detailed quality assessments. It is not supported by empirical evidence, it can be time-consuming and it is potentially misleading.