Evidence Synthesis
Number 42

Screening for Developmental Dysplasia of the Hip

U.S. Department of Health and Human Services
Agency for Healthcare Research and Quality
www.ahrq.gov

This report may be used, in whole or in part, as the basis for development of clinical practice guidelines and other quality enhancement tools, or a basis for reimbursement and coverage policies. AHRQ or U.S. Department of Health and Human Services endorsement of such derivative products may not be stated or implied.

AHRQ is the lead Federal agency charged with supporting research designed to improve the quality of health care, reduce its cost, address patient safety and medical errors, and broaden access to essential services. AHRQ sponsors and conducts research that provides evidence-based information on health care outcomes; quality; and cost, use, and access. The information helps health care decisionmakers—patients and clinicians, health system leaders, and policymakers—make more informed decisions and improve the quality of health care services.

Evidence Synthesis
Number 42

Screening for Developmental Dysplasia of the Hip

Prepared for:
Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
540 Gaither Road
Rockville, MD 20850
http://www.ahrq.gov

Contract No. 290-02-0024
Task No. 2
Technical Support of the U.S. Preventive Services Task Force

Prepared by:
Oregon Evidence-based Practice Center
Oregon Health and Science University
3181 SW Sam Jackson Park Road
Portland, Oregon 97239

Scott Shipman, MD, MPH, Principal Investigator
Mark Helfand, MD, MPH, Co-Investigator
Peggy Nygren, MA, Research Associate
Christina Bougatsos, Research Assistant

March 2006

Preface

The agency for Healthcare Research and Quality (AHRQ) sponsors the development of Evidence Syntheses through its Evidence-based Practice Program. With guidance from the U.S. Preventive Services Task Force (USPSTF)* and input from Federal partners and primary care specialty societies, the Evidence-based Practice Center at the Oregon Health & Science University systematically reviews the evidence of the effectiveness of a wide range of clinical preventive services, including screening, counseling, and chemoprevention, in the primary care setting. The Evidence Syntheses—comprehensive reviews of the scientific evidence on the effectiveness of particular clinical preventive services—serve as the foundation for the recommendations of the USPSTF, which provide age- and risk-factor-specific recommendations for the delivery of these services in the primary care setting. Details of the process of identifying and evaluating relevant scientific evidence are described in the "Methods" section of each Evidence Synthesis.

The evidence Syntheses document the evidence regarding the benefits, limitations, and cost-effectiveness of a broad range of clinical preventive services and will help further awareness, delivery, and coverage of preventive care as an integral part of quality primary health care.

AHRQ disseminates the Evidence Syntheses on the AHRQ Web site (http://www.ahrq.gov/clinic/uspstfix.htm) along with recommendations of the USPSTF. The Recommendation Statements are also available through the National Guideline Clearinghouse (http://www.ngc.gov).

We welcome written comments on this Evidence Synthesis. Comments may be sent to: Director, Center fro Primary Care, Prevention, and Clinical Partnerships, Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, MD 20850.

Carolyn Clancy, M.D.
Director
Agency for Healthcare Research and Quality

Helen Burstin, M.D., M.P.H.
Director
Center fro Primary Care, Prevention, and Clinical Partnerships

*The USPSTF is an independent panel of experts in primary care and prevention first convened by the U.S. Public Health Service in 1984. The USPSTF systematically reviews the evidence on the effectiveness of providing clinical preventive services—including screening, counseling, and chemoprevention—in the primary care setting. AHRQ convened the USPSTF in November 1998 to update existing Task Force recommendations and to address new topics.

Acknowledgements

This evidence synthesis was funded by the Agency for Healthcare Research & Quality (AHRQ) for the U.S. Preventive Services Task Force (USPSTF). The investigators thank, Ronald Turker, MD, for serving as a consultant, the USPSTF liaisons, Carol Loveland-Cherry, PhD, RN, Leon Gordis, MD, MPH, DrPH, Virginia Moyer, MD, MPH, and Barbara Yawn, MD, MSc, the AHRQ Medical Officer, Iris Mabry, MD, expert reviewers listed in the Appendix of this report for commenting on draft versions, Andrew Hamilton, MLS, MS for conducting the literature searches, and Miranda Walker for assisting with the manuscript.

Requests for Reprints

Reprints are available from the AHRQ Web site at www.ahrq.gov (click on Preventive Services).

Disclaimer

The authors of this article are responsible for its contents, including any clinical or treatment recommendations. No statement in this article should be construed as an official position of the Agency for Healthcare Research and Quality or the U.S. Department of Health and Human Services.

Contents
Screening for Developmental Dysplasia of the Hip

Figures

Tables

Appendices

Structured Abstract

Context: Developmental dysplasia of the hip (DDH) can lead to the later development of chronic pain, osteoarthritis, and limitations in activity. Screening for DDH has been practiced for over 40 years, but recommendations from major professional societies differ.

Objective: To synthesize the evidence on risks and benefits of screening for DDH.

Data Sources: MEDLINE (through Sept, 2004), Cochrane CENTRAL, and previous comprehensive literature reviews.

Study Selection: We focused our review on information gaps identified in previous reviews conducted for the American Academy of Pediatrics and the Canadian Task Force on Preventive Health Care. Specifically, we focused on comparative studies of clinical examination vs. ultrasound screening; studies of the effect of nonsurgical and surgical treatments for DDH on functional outcomes; and studies reporting rates of avascular necrosis with different interventions.

Data Extraction: Using present criteria, the authors assessed the quality of included trials and abstracted information about settings, patients, interventions, and outcomes.

Data Synthesis: No published trials directly link screening to improved functional outcomes. Clinical examination and ultrasound identify somewhat different groups of newborns at risk for DDH; the lack of an untreated cohort or definitive gold standard made it impossible to estimate sensitivity and specificity for the different tests. Few studies examine the functional outcomes of patients who have undergone therapy for DDH. Due to the high rate and unpredictable nature of spontaneous resolution of DDH

and the absence of comparative studies of intervention vs. no intervention, the effectiveness of interventions is not known. Avascular necrosis (AVN) of the hip, the most common and most severe harm of all treatments for DDH, can result in growth arrest of the hip and eventual joint destruction with significant disability. Reported rates of AVN very widely.

Conclusion: Screening with clinical examination or ultrasound can identify newborns at risk for DDH, but due to the high rate of spontaneous resolution of neonatal hip instability and dysplasia and the lack of evidence of the effectiveness of interventions on functional outcomes, the net benefits of screening are not clear.

Key Words: DDH, Hip Dysplasia, mass screening

Chapter 1. Introduction

Developmental dislocation of the hip can lead to premature degenerative joint disease, impaired walking, and pain. Surgery is often necessary once these complications have occurred. Hip instability can be treated nonsurgically if it is detected early. Neonatal screening, which has been practiced for almost four decades, is intended to reduce the need for surgery, prevent degenerative joint disease, pain, and mobility limitations.

This evidence synthesis focuses on screening and intervention for developmental dysplasia of the hip (DDH) in physiologically normal infants from birth through 6 months. The review was conducted for the U.S. Preventive Services Task Force (USPSTF), which had no previous recommendations for this condition. Two systematic reviews of DDH have been published, one by the Canadian Task Force on Preventive Health Care (CTFPHC)[1] and another by the American Academy of Pediatrics (AAP).[2, 3] This evidence synthesis will summarize this previous work with a focus on how methods and conclusions agree and differ, and incorporates published studies since these reviews were completed.

Burden of Condition

DDH represents a spectrum of anatomical abnormalities in which the femoral head and the acetabulum are in improper alignment and/or grow abnormally. The precise definition of DDH is controversial.[2, 4] The spectrum includes hips that are dysplastic, subluxated, dislocatable and dislocated. Clinical instability of the hip is the traditional hallmark of the disorder. In an unstable hip, the femoral head and acetabulum may not

have a normal tight, concentric anatomic relationship, which can lead to abnormal growth of the hip joint and may result in permanent disability.

Nonspecific instability in the hip is a common finding in newborns.[5] This is particularly true in females, in whom the maternal hormone relaxin may contribute to ligamentous laxity. More than 80% of clinically unstable hips noted at birth have been shown to resolve spontaneously.[6] However, due to the potential for subsequent impairment and the widespread belief that earlier treatment leads to improved outcomes, screening newborns for DDH has become commonplace.

Estimates of the incidence of DDH in infants vary between 1.5 and 20 per 1000 births.[1] The incidence of DDH in infants is influenced by a number of factors, including diagnostic criteria, gender, genetic and racial factors, and age of the population in question.[3] The reported incidence has increased dramatically since the advent of clinical and sonographic screening, suggesting possible overdiagnosis.[4] Risk factors for the development of DDH include gender, family history of DDH, breech intrauterine positioning, and additional in utero postural deformities.[7-9] However, the majority of cases of DDH have no identifiable risk factors.[10]

The most common methods of screening for DDH involve the physical examination of the hips and lower extremities. Provocative testing includes the Barlow and Ortolani procedures, which involve adduction of the flexed hip with gentle posterior force, and abduction of the flexed hip with gentle anterior force, respectively. The Barlow test attempts to identify a dislocatable hip,[6, 11] while the Ortolani exam attempts to relocate a dislocated hip.[12] Additional findings reported on physical examination in infants include asymmetry of gluteal and thigh skin folds, discrepant leg lengths, and diminished range

of motion (particularly abduction) in an affected hip.[3] Due to variations in technique, the provocative Barlow and Ortolani tests have been shown to have a high degree of operator dependence.[13] In addition, confusion about the identification of a "click" versus a "clunk" on these tests, and the significance of each of these findings, can lead to disparate conclusions between examiners.

Ultrasonography and radiography are also used to screen for DDH. X-ray is less accurate in the first 3-4 months of life, when the bones of the hip are not completely ossified. The use of ultrasonography and/or radiography in screening has been controversial, particularly due to reports of high false positive rates leading to unnecessary and potentially harmful follow-up and intervention.[14] Despite the controversy, ultrasound has been widely incorporated into DDH screening programs in many developed countries.[15, 16]

Healthcare Interventions

Intervention for DDH includes both nonsurgical and surgical options. A variety of abduction devices are used to treat DDH nonsurgically, with the Pavlik method among the most common. These devices place the legs and hips in an abducted and flexed position in an effort to promote stabilization of the hip joint. The duration of treatment varies from center to center. Complications of nonsurgical therapy are not trivial, with avascular necrosis of the femoral head among the most serious.[1]

Surgical intervention is necessary when DDH is severe, when it is diagnosed late, or after an unsuccessful trial of nonsurgical methods.[17] Many surgical procedures are used to treat DDH (Appendix 1). Most involve reduction of the femoral head into the

acetabulum, with or without additional procedures on the adductor tendons, the femur, or the acetabulum. Preoperative management often includes a period of traction, and postoperative management typically includes a period of fixed positioning in a spica cast. The duration and specific approach to pre- and post-operative management are highly variable. Surgical intervention places the hip at risk of avascular necrosis, in addition to standard operative risks including general anesthesia, intraoperative complications, and post-operative wound infections.

For the purposes of this review, all nonsurgical abduction therapy was considered as a whole, distinct from all surgical procedures which were also considered collectively. Since closed manual reduction of the hip typically requires general anesthesia, it was considered along with other surgical procedures.

Prior Recommendations

In 2000, the AAP used a combination of expert panel, decision modeling, and evidence synthesis to develop DDH screening guidelines. The AAP recommended universal screening of newborns by serial physical examination, with 2 week follow-up examination for equivocal findings and referral to an orthopedist for positive Barlow and Ortolani tests.[2, 3] They also emphasized the importance of considering risk factors in the approach to screening, recommending ultrasound in females born breech, and recognizing ultrasound of all infants born breech as a reasonable approach. However, they did not recommend universal ultrasound screening. The AAP report did not examine the effectiveness of therapy (Table 1).

In 2001, the CTFPHC also recommended universal screening via serial physical examination of the hips until the patient is walking.[1] They recommended against the use of ultrasound or radiography in a selective approach to screening, in contrast to the AAP report, and concurred with the AAP report in opposing universal ultrasound screening. The CTFPHC examined the literature on abduction therapy for DDH, and concluded that there was insufficient evidence to evaluate the effectiveness of this intervention (Table 1).

Scope of Evidence Synthesis

The analytic framework (Figure 1) and key questions (Figure 2) guiding the literature review were developed in consultation with liaisons from the USPSTF. We focused on screening in infants from birth through 6 months of age. We excluded so-called teratological DDH, that occurring in children with neuromuscular disorders or other congenital malformations. We included literature on the effects of both nonsurgical (abduction braces) and surgical interventions on functional outcomes, including gait, pain, physical functioning, activity level, peer relations, family relations, school and occupational performance.

The key questions examine critical links in the logic underlying screening. To be effective, screening must identify cases of DDH earlier than they would be identified in the usual course of care (Key Questions 2, 3). In addition, early identification must lead to earlier treatment, and earlier treatment must lead to better functional outcomes than late treatment (KQ5). Finally, the benefits of early identification and treatment must outweigh the harms of screening and of the treatments themselves (KQ4, 6).

Chapter 2. Methods

Literature Search Strategy

The most recent systematic reviews of screening for DDH, by the AAP and the CTFPHC, targeted many of the same questions as this report. We analyzed their reviews to focus the search strategy and eligibility criteria for our review. When questions had substantial overlap, we reviewed all studies identified in these reviews and searched the literature for studies published subsequently (after 1996 for the AAP review and 2000 for the CTFPHC review).

For most key questions, relevant studies were identified from multiple searches of MEDLINE (1966 to January 2005) and the Cochrane Library databases through June of 2004. Search strategies are described in Appendix 2. Additional articles were obtained by reviewing reference lists of other pertinent studies, reviews, editorials, and websites, and by consulting experts. We modified this strategy after reviewing the two previous systematic reviews (see Results section, subsection Previous Systematic Reviews). Specifically, for assessments of screening modalities in Key Question 3, we examined the literature beginning in 1996, the year in which the AAP review concluded.

Inclusion/Exclusion Criteria

Investigators reviewed all abstracts identified in the searches and the previous systematic reviews and determined eligibility by applying inclusion and exclusion criteria specific to key questions (Appendix 3). Full-text papers of included abstracts were then reviewed for relevance. Eligible studies had English-language abstracts, were applicable

to U.S. clinical practice, and provided primary data relevant to key questions. Initial screening had to be done in children less than 6 months of age, and screening studies needed to be prospective, primary care based or population based in design. Studies of risk factors also had to be primary care based or population based. Intervention and outcomes studies had to report results of children diagnosed before 6 months of age, and interventions had to be employed earlier than 1 year of age on average. For intervention studies, we were particularly interested in functional outcomes, including: gait, pain, physical functioning, activity level, peer relations, family relations, school and occupational performance. For noninvasive interventions, another potential benefit is a reduced need for surgery later in childhood. Therefore, intervention studies were eligible if they reported one of these functional outcomes and/or a subsequent need for surgery. We excluded studies that reported only radiological reports of anatomic structural relationships and development, which have not been shown to be valid predictors of functional outcomes. For avascular necrosis (AVN), the predominant harm from interventions, studies needed to report the rate of this complication in the treated patient population, meet age-based inclusion criteria, have at least 1 year of follow-up, and not experience excessive (>50%) loss to follow-up.

We used a "best evidence" approach[18]; that is, for each key question, we included studies with weaker designs only if better-designed studies were not available. Case reports, series with 5 or fewer subjects, editorials, letters, nonsystematic review articles, and commentaries were also excluded from the evidence review.

Data Extraction and Synthesis

Data were extracted from each study, entered into evidence tables, and summarized by descriptive and statistical methods as appropriate. We rated the internal validity of each included study using criteria specific to different study designs developed by the USPSTF (Appendix 4).[19] The USPSTF quality criteria can be used to appraise controlled trials, observational, comparative studies such as cohort and case-control studies, and studies evaluating the performance of diagnostic tests. Studies with flaws deemed to invalidate the results were labeled as poor in quality, and were not included in the evidence report.

Most studies of DDH are observational, uncontrolled or poorly controlled, and have serious flaws in design (grade of II-3 or III according to the original USPSTF classification.) There are no USPSTF criteria to rate such studies good, fair, or poor, but we highlight their limitations. To assess the quality of these studies, we considered the following: study design, clarity of diagnostic standards, comparability of subjects, variation in screening approach and/or intervention protocol, duration of follow-up, loss to follow-up, efforts to control for confounding and minimize bias, masking of outcome assessors, and validity and standardization of outcomes measured.[20]

Size of Literature Reviewed

Investigators reviewed 1,145 abstracts of English-language articles identified by the searches, excluding 679 citations on first review (Appendix 5). A total of 466 full-text articles were retrieved and reviewed; 416 were from the electronic searches and 50 were from reference lists or experts' suggestions (expert reviewers listed in Appendix 6).

Thirteen papers about risk factors; 59 about screening, including 3 controlled trials; 5 about harms of screening; 47 about interventions and harms of interventions, including no controlled trials; and 8 about cost met the inclusion criteria. Review of an additional 544 abstracts of non-English language articles did not identify any additional controlled trials.

Chapter 3. Results

Previous Systematic Reviews

The AAP recommendations were based on an extensive review of the literature, including Medline and EMBASE searches through June, 1996.[2, 3] Articles were included in the review if they helped to estimate one or more probabilities in a decision model comparing five screening strategies: no screening, screening high-risk newborns by physical examination alone; screening all newborns by physical examination alone; screening all newborns with ultrasound; and screening all newborns by physical examination conducted by an orthopedic surgeon. A total of 118 articles (5 comparative trials and 113 observational studies) were included in the review. The authors noted that no evidence was available for 13 of the 30 probabilities they sought to estimate.

The AAP review methods differed from ours in several respects. First, they used a different system to assess the quality of individual studies. Specifically, they developed a 7-item, 21-point quality scale. One item graded the method of assignment to groups (that is, “random”=3 points, “comparative arm”=2 points, “single arm”=1 point, and “haphazard”=0). Other scale items rated the degree to which the study results were applicable to one or more parameters in the decision model. By contrast, the USPSTF

rating system examines characteristics of the study related to the internal validity of the results. Second, the AAP model incorporated experts' opinions when there were gaps in the published evidence. Thus, the quality of evidence supporting the reports' findings is quite variable.

The AAP review used 106 observational studies (which they described as "case series") to estimate the chance of a positive screening examination for different patient populations and with different screening modalities. For example, they used 48 observational studies published between 1956 and 1996 to estimate the probability of a positive physical examination when screening was conducted by pediatricians. After examining the articles included in the AAP review, we determined that 36 of the 48 studies of screening by clinical examination reported results of screening in a population relevant to our review. We concluded that the AAP report made valid estimates of the rates of positive clinical screening examinations through 1996.

The AAP review found limited evidence on the yield of universal ultrasound and the value of serial examinations for DDH. More over, the AAP report did not focus on the comparative yield of clinical examination and ultrasound when both are applied to the same population. Also, while it examined how well risk factors predict a positive screening test, it did not examine how well risk factors predict confirmed cases of DDH.

Literature examining the effectiveness of nonsurgical or surgical interventions was outside the scope of the AAP review; assumptions about the effectiveness of these interventions were based on expert opinion. Its review of rates of AVN, which focused on the relation between the risk of AVN and the age of referral, identified fair-to-good quality evidence.

The CTFPHC report[1] on DDH sought to answer many of the questions identified in the present review. This report was not accompanied by a comprehensive technical report as was the AAP study. The CTFPHC report cites fair evidence supporting serial clinical screening examination, but upon further review the evidence cited is sparse (see KQ 1). Their review of the role of ultrasound in screening focused on the single available controlled trial,[21] but also summarized findings from 32 additional studies, predominantly descriptive in nature. The CTFPHC review also examined the nonsurgical intervention literature, but their criteria for evaluating the intervention literature were not explicit; their review included studies with radiological (rather than functional) outcomes. They found insufficient evidence to assess the effectiveness of abduction therapy. Finally, they concluded that a period of supervised observation is warranted prior to initiating therapy in hips diagnosed with DDH at birth, given the high rate of spontaneous resolution. Appendix 7 compares the degree to which the literature in the CTFPHC report met our inclusion criteria.

Key Question 1. Does Screening for DDH Lead to Improved Outcomes (including reduced need for surgery and improved functional outcomes such as: gait, physical functioning, activity level, peer relations, family relations, school and occupational performance)?

There are no prospective studies—either randomized or observational—comparing a screened to a non-screened population with measurement of functional outcomes after an

adequate period of follow-up. There are also no controlled trials that compare surgical or nonsurgical treatment for early DDH to observation only.

In theory, early application of noninvasive treatments (e.g., a harness) to obtain a concentric and stable reduction of the femoral head in the acetabulum may obviate the need for surgery later on. However, the evidence that screening leads to a reduced rate of surgery is weak and indirect. The 2000 CTFPHC report, citing several descriptive studies, concluded "With serial clinical examination, the operative rate for DDH has decreased by more than 50% to 0.2-0.7% per 1000."[1] It should be noted that this reduction was observed at an ecological level: descriptive studies in screened populations were compared, indirectly, to unscreened populations or to historical rates. The studies were not comparative and did not report functional outcomes. In addition, while some studies suggest that surgical rates have declined since the adoption of universal screening programs, they do not indicate why. The decline might be attributable to increased rates of screening, but other factors, such as wider use of a period of observation before recommending surgery, could also account for the declining use of these surgical procedures.

The measure used in many comparative studies was the proportion of infants and children with DDH who had surgical intervention. If screening identifies more cases than usual care, it could reduce this proportion even if the same number of cases required surgery as before. For this reason it is difficult to determine whether a decrease in the surgical rate over time reflects the efficacy of noninvasive intervention or the inclusion of additional cases in the denominator who are at little or no risk of requiring surgery.

The findings are also inconsistent: some studies observed a decrease in operative rates,[22-25] while others saw no change[26, 27] or an increase.[28-30] Ascertainment of cases was often flawed, and the studies span several decades, making it difficult to assess whether the varied results represent artifacts of data quality, secular trends, or differences in local practice styles.[31] These studies are also limited because they typically do not follow the screen-negative population with the same vigilance as the screen positive population, and experience significant loss to follow-up in the screen positive population that can bias the outcomes.

More recent studies also have conflicting results. In 1998, the MRC Working Party on Congenital Dislocation of the Hip reported operative rates in a randomly selected, population-based survey of 20% of all births in the U.K.[31] After adjustment for differences in ascertainment that had been overlooked in previous reports, the incidence of a first operative procedure for congenital dislocation of the hip was similar before and after screening was introduced (pre-screening rate range 0.66 – 0.85 per 1000, post-screening rate 0.78 per 1000 live births, 95% CI 0.72-0.84). Even in the screening era, 70% of the cases reported by surgeons to the registry had not been detected by screening. In 1999, Australian investigators reported the operative rate in the post-screening era using an existing perinatal database with information about birth defects and an inpatient discharge database to identify infants with congenital dislocation of the hip.[32] In contrast to the U.K. study above, they reported an operative rate of 0.46 per 1000 live births and found that 97.6% of congenital dislocation cases were diagnosed before 3 months of age. The causes behind conflicting findings such as in these two studies are unknown.

Key Question 2. Can Infants at High Risk for DDH be Identified, and Does This Group Warrant a Different Approach to Screening than Children at Average Risk?

Risk factors are considered an adjunct to, rather than a substitute for, universal screening by physical examination. For example, the AAP recommends using risk factors to identify newborns whose risk for DDH may exceed the comfort level of physicians, prompting additional screening using ultrasound. The rationale for this approach is that, in high-risk newborns, clinical examination alone will miss many cases of DDH that ultrasound can identify. The assumptions underlying this approach are (1) risk factors can identify a group of newborns at a high risk of DDH and (2) ultrasound is more sensitive than clinical examination for identifying infants at risk of complications from DDH.

In case control and observational studies, breech positioning at delivery, family history of DDH, and female gender have been most consistently shown to have an association with the diagnosis of DDH. Additional risk factors may include maternal primiparity, high birthweight, oligohydramnios, and congenital anomalies.

Lehmann and colleagues conducted a meta-analysis of studies published through 1996 to estimate the probability of having a positive screening test for the three leading risk factors.[2] Breech females (84/1000) had a dramatically higher than average risk (8.6/1000 for all newborns) of being screen-positive, followed by family history positive females (24/1000), breech males (18/1000), females with no risk factors (14/1000), and males with no risk factors having the lowest risk (3/1000).

The DDH reference standard in their synthesis was a positive Barlow and Ortolani test at the newborn screening examination. While this is a commonly used and reasonable measure of the disorder, it may overestimate the number of infants requiring therapy. Primary care and population-based cohort studies[33-43] that included one or more of the major risk factors are summarized in Table 2. Consistently, only a minority (10-27%) of all infants diagnosed with DDH in population-based studies have identified risk factors (with the exception of female gender)[37, 39, 40, 42] and among those with risk factors, between 1% and 10% have DDH.[37, 40, 42] This wide range illustrates the impact of the reference standard on the relative importance of risk factors. Those studies with a stricter standard for diagnosing "true" DDH, for instance limited to those cases that receive treatment, demonstrate substantially lower rates of DDH among those with risk factors. For example, a recent cohort study of 29,323 births at one hospital, the prevalence of treated DDH was 20/1000 in breech females (vs. 110/1000 based upon the clinical exam), 12/1000 in family history positive females, 4/1000 in breech males, 5/1000 and 0.3/1000 in females and males with no risk factors, respectively.[35] The substantial differences (4 fold in the case of breech females) in prevalence between the AAP estimates and this study reflect different diagnostic standards, and impact the predictive value of risk factors for DDH. More conservative estimates based upon "true" DDH makes the value of routine ultrasound for patients with given risk factors less certain. From a primary care perspective, a prospective, practice-based cohort study of a risk scoring or other risk assessment tool would provide the strongest evidence about the yield of selective screening of high-risk infants.

Several potential biases should be considered in evaluating risk factor data. In studies where the examiner is aware of patients' risk factor status, the diagnosis of DDH may be overestimated due to more careful or thorough examinations or more aggressive follow-up and reexamination in infants with known risk factors. Moreover, in retrospective studies researchers apply criteria to improve the reliability of their record review; this approach, while necessary to conduct such a study, reduces the influence of an equivocal or inaccurate history. A predictor such as family history may be less reliable in a prospective, practice-based study than in case control studies which exclude patients (charts) that have equivocal or incomplete information about it. Finally, investigators' awareness of the subjects' final diagnoses could have influenced the way they handled risk factor information.

Key Question 3. Does Screening for DDH Lead to Early Identification of Children with DDH?

Clinical screening for DDH includes the provocative Barlow and Ortolani tests of hip stability, and assessment of range of motion of the hip in abduction. In addition to clinical examination, the approach to screening may include imaging of the hip, traditionally by radiography and more typically today by ultrasound. Ultrasound methods include both static and dynamic assessments of the hip, and its use varies widely across developed nations. All methods used to screen for DDH are variably subjective and operator-dependent.

Recent prospective population-based and primary care practice-based studies[14, 16, 35, 44-47] offering a within-group comparison of clinical examination and ultrasound screening

are summarized in Table 3. Randomized trials[21, 48, 49] of different screening modalities are summarized in Tables 4 and 5.

KQ 3a. What is the accuracy of clinical examination and ultrasound? To measure sensitivity directly in a prospective study, infants who had negative initial screening tests must be followed and examined at older ages to identify false negative initial test results. Measuring sensitivity is also difficult because results of the Barlow test can be classified into several levels, rather than just two ("positive" or "negative"). Conversely, measuring specificity and false positives is difficult because, in most studies, all infants who have a positive screening test are treated with a nonsurgical intervention; the great majority improve, and it is impossible to say how many of them "responded" and how many of them did not have DDH in the first place.

Assessing the impact of a screening program on the rate of late diagnosis of DDH provides an indirect measure of sensitivity. It is apparent that screening tests performed soon after birth identify some individuals at risk of developing DDH sooner than they would otherwise be identified: most children would otherwise not come to medical attention until the age of walking (approximately 1 year) in most cases. However, it is difficult to quantify the impact of screening tests on the incidence of late diagnosis with the available literature. Studies of the impact of screening programs on the frequency of late diagnosis have had mixed results.[23-25, 28, 32, 50-62] Most of these studies report the experience of a screening program in a defined geographic or hospital service area over many years. The comparisons are ecological, and these studies have the same methodological problems as those that examined the effect of screening on rates of

surgical treatment (discussed above, Key Question 1). Some studies in this group reported that, after a screening program was adopted, late diagnosis was very rare, while others report that screening had no effect on the rate of late diagnosis, and that unexplained fluctuations in late diagnosis rates were observed from year to year within the post-screening era (Figure 3 and Figure 4).[21, 23-25, 27-29, 36, 40, 50, 55, 60]

The lack of a practical confirmatory "gold standard" diagnostic test for DDH makes it difficult to assess—or define—false positives. Various reference standards appear in the literature, including positive clinical examination, ultrasound confirmation, radiographic confirmation, arthrography, persistence of abnormal findings on serial exam or ultrasound over weeks to months, diagnosis by an orthopedist, and use of treatment. The most meaningful reference standard defines "true" DDH as "those neonatal hips, which, if left untreated, would develop any kind of dysplasia and, therefore, are to be included in the determination of DDH incidence."[4]

To apply this standard, a cohort study must follow infants for a long enough period without applying any treatment, in order to determine whether or not the abnormal findings persist and lead to clinical problems. In one good-quality prospective cohort study that followed untreated infants for 2 to 6 weeks, approximately 9 of 10 infants with initially abnormal ultrasound examinations revert to normal.[4] Similarly, by 2 - 4 weeks of age, over 60% of infants identified at birth by abnormal clinical examination (Barlow or Ortolani tests) have reverted to normal when judged by repeat clinical examination or by ultrasound examination.[6, 11, 63] Longer prospective studies[21, 35, 63-68] and a systematic review of observational studies of ultrasound screening[69] demonstrate that in untreated

hips, mild dysplasia without frank instability usually (consistently over 90%) resolves spontaneously between 6 weeks and 6 months.

Table 3 includes population-based (or primary care clinic based) cohorts screened by clinical examination as well as ultrasound screening, published since the 1996 endpoint of the AAP review.[14, 16, 35, 44-47] Despite variation in the reference standards used in these studies, several important findings emerge. First, a high proportion of hips diagnosed with minor findings of dysplasia undergo spontaneous resolution**.** It is important to note that minor dysplasia is not identified by clinical exam, but only by ultrasound. Due to the identification of anatomic variations that are marginal and self-limited, the potential exists for over-treatment on the basis of ultrasound. On the other hand, in 4 of the 7 studies in Table 3, 38% - 87% of abnormal findings on clinical exam were not DDH, leading to a similar risk of unnecessary therapy on the basis of clinical examination.[16, 44, 45, 47] Very few of these studies followed patients longitudinally, particularly those patients who did not screen positive by exam or ultrasound.

In the first 4-6 months of life, ultrasound has been deemed to be a more appropriate test than radiographs for anatomic hip abnormalities as well as instability of the hip, due to incomplete ossification of the femoral head in early infancy. Though no study addressed the comparative value of ultrasound to radiograph in the 4-6 month time-frame, there is strong endorsement of this approach in the literature, ranging from historical studies reporting on timing of ossification and analyzing the technical challenges of hip radiography in the young infant,[70, 71] to contemporary systematic reviews.[2, 3]

However, ultrasound screening is not without its shortcomings. In addition to the high rate of identification of nonpathological hip findings summarized above, the most widely used ultrasound-grading system, Graf classification,[72] has come under scrutiny. The Graf score is used in the vast majority of the screening literature to differentiate normal hips from immature hips from minor dysplasia from major dysplasia; and stable from unstable, subluxable, and dislocatable/dislocated. Many studies base treatment decisions on these classifications. A study examining the reliability of Graf classification found that, among normal hips, intra- and inter-observer reliability is quite high, with a 98% chance of having the same assessment on future readings. However, among ultrasounds read as abnormal by at least one person, intra-observer reliability was moderate (kappa = 0.41) but inter-observer reliability was fair (kappa = 0.28). In addition, knowledge of the patients' history and physical exam vs. blinded review of the ultrasound lowered the intra-observer kappa from 0.41 to 0.37.[73]

Another study found moderate agreement between observers on determining morphology by subjective reading (kappa = 0.5), but this decreased to 0.3 when objective measurements of anatomic relationships were conducted. Grading of stability was moderate (kappa 0.42) between observers, when dislocatable and dislocated hips were grouped together. This study estimated that the decision to treat would have been affected in 2.4% of cases due to discordance between reviewers.[74] Considerable effort had been given to standardizing ultrasound assessment in this study, including a training session and 100 repetitions of conducting measurements before the start of the study. Still another study found ultrasound reliability to be similarly suspect, with kappas ranging from 0.52 -0.68 and 0.09 to 0.30 for intraobserver and interobserver agreement,

respectively, across seven anatomic measures used in grading DDH.[75] These findings raise concerns about the operator dependence of this evaluation for DDH, and may shed light on the variability of ultrasound screen positive rates found in the literature.

While there are no trials or comparative studies of a screened to an unscreened population, 2 randomized controlled trials[48, 49] and 1 nonrandomized controlled trial[21] provide some insight into the accuracy of clinical examination. These trials reported data about test performance of one screening strategy versus another (Table 5). The first randomized controlled trial (RCT) compared universal ultrasound screening to selective screening at a population level.[49] In the trial, patients at the University of Trondheim, Norway were randomized over a 5 year period to one of two groups: clinical exam and ultrasound or clinical exam and selective ultrasound. In the first group, each of the 7840 patients received clinical exam and ultrasound. In the other group, 7689 received clinical exam alone or, if they had risk factors (abnormal exam, breech, family history, foot deformities), ultrasound and clinical exam. In the selective ultrasound group, 5 infants presented between 5-6 months with previously undiagnosed DDH, whereas in the universal screening group there was only 1 case of late diagnosis. In all these late-presenting cases, treatment with an abduction brace was implemented and the hips were reported to be normal upon follow-up. Overall treatment rates were equivalent in the two groups.

The second RCT[48] included 629 patients who had been diagnosed with unstable hips on screening examination and were referred to 33 specialty centers in the United Kingdom (UK). The subjects were randomized within the specialty centers to receive ultrasonographic hip examination (n=314) or clinical assessment alone (n=315). A total

of 90% of patients in the ultrasound group received an ultrasound in the first 8 weeks of life; 8% in the no-ultrasound group received an ultrasound. Compared to those in the ultrasound group, infants in the no-ultrasound group were treated more often (50% vs. 40%) and earlier (98/150 vs. 42/117 treated in the first 2 weeks of life). The need for surgical treatment (8% vs. 7%), age at surgical treatment (31 vs. 29 weeks), mean number of visits at outpatient clinics (4 in each), total hip-related hospitalizations (30 vs. 23) and the occurrence of definite or suspected avascular necrosis (5 vs. 8) were not significantly different between the two groups. Thus, despite a higher rate and earlier initiation of treatment in the clinical examination only group, the non-functional "outcomes" of the two groups were quite similar. This suggests that, in the specialty setting, clinical examination alone may lead to a greater degree of unnecessary treatment than that which occurs when an abnormal clinical examination is followed up with evaluation by ultrasound.

An earlier controlled trial, conducted in 1994, compared 3613 infants in a universal screening program to 4388 in a selective screening program, and 3924 who received only clinical examination.[21] In the selective ultrasound cohort, a positive clinical examination was considered to be a risk factor prompting ultrasound. The study concluded that universal ultrasound had a significantly higher treatment rate overall, but no higher rate among high risk infants. There was a nonsignificant trend toward a lower rate of cases diagnosed after 1 month of age in the universal screening patients. Among those not treated, many more children with mildly dysplastic hips were identified by ultrasound, resulting in more follow-up visits and ultrasounds for a greater number of patients in the universal screening approach.

b) How does the age of the child affect screening parameters? Irrespective of reference standard, the clinical exam approach to diagnosis for DDH shifts over time. Barlow and Ortolani tests become less sensitive as infants age, due to factors including increased strength, bulk, and size (Key Question 3b).[1, 3] In their place, assessment of hip abduction becomes the preferred examination, because infants with dislocated hips have increased contractures of the hip adductors.[3] Specificity of examination improves as infants' age, because the hips of the newborn infant are more likely to exhibit transient and clinically insignificant laxity than they will subsequently.[11] Two recent studies provide indirect insight into the changing signs of DDH as the infant ages. In a study of 1071 referred infants at one center, only 2 of 34 (6%) hips in patients with positive Barlow or Ortolani tests, confirmed as dislocatable by ultrasound, had any limitation in abduction in patients at 1-2 weeks of age, suggesting poor sensitivity in newborns.[76] Specificity of limited hip abduction in newborns was also poor. Among 203 1-2 week old infants with limited abduction, <20% had abnormalities on ultrasound. These findings contrasted with older children: of the eight patients who presented after six months of age with dislocatable hips, hip abduction was limited in 7 (87.5%). The second study, a prospective observational study limited to infants greater than 3 months of age (N=683), found that unilateral limited hip abduction had a sensitivity of 69% (156/226), and a specificity of 54% (247/457).[77] The reference standard in this study was any ultrasound abnormality; among subluxable and dislocatable hips, sensitivity of limited hip abduction was > 82%. Of the patients with limited abduction and normal ultrasound findings (N=136), none showed any abnormalities on examination, and all walked normally without a limp at 5 years of age. Though not conclusive, these studies

suggest that hip abduction is a relatively insensitive and nonspecific marker of DDH in early infancy, but becomes more accurate after 3-6 months of age and with more severely affected hips.

Additional physical examination findings sometimes linked to DDH include asymmetrical gluteal and thigh skinfolds, and leg length discrepancy. No studies from the past 40 years were identified which assessed the value of these findings in diagnosing DDH. Barlow pointed out the lack of utility of asymmetric skin folds due to their poor sensitivity and specificity,[6] and Palmén studied 500 random newborns, finding that 27% had no thigh skinfolds, 40% were symmetrical, and 33% asymmetrical; 4 of these 500 babies had an abnormal provocative test of stability, of which 2 had symmetrical skinfolds.[70]

3c) How does the educational level and training of the screener impact screening? The degree of training and experience with the clinical examination of the hip in infants has been shown to be a strong predictor of the test characteristics (Key Question 3c). Pediatricians have been shown to have a case identification rate of 8/1000, whereas orthopedists identify approximately 11/1000.[2] In one single site longitudinal study, during periods when the number of pediatricians involved in the screening program increased (holding steady the number of newborns screened), a greater number of cases of DDH were missed despite an increased rate of suspected cases identified.[78] This finding may suggest that screening accuracy suffers when an examiner has less ongoing experience in the exam technique. Two studies show that having duplicate blinded examinations by a pediatrician and an orthopedist improves the sensitivity, specificity, and predictive value of clinical exam screening.[79, 80] Additional studies show that well-

trained non-physicians, including physiotherapists and neonatal nurse practitioners, perform at least as well as physician examiners, and better than physician trainees.[81-83]

In several studies comparing pediatricians with orthopedic surgeons, the surgeons review a subset of hips found to be positive or questionable by the previous examiner. This may happen days after the initial examination. Also, the surgeons often have at their disposal the results of ultrasonography, and their clinical examination is not blinded from the ultrasound exam. Not surprisingly, such studies show a higher sensitivity and specificity of clinical examination in the hands of the specialist.

Key Question 4. What Are the Adverse Effects of Screening?

Dislocation. While it has been suggested that the examination of already-lax newborn hips might cause injury or dislocation,[84] we identified little research that sought to test this hypothesis. Three studies provides some insights[85-87] An autopsy study examined 10 hips in stillborn infants, 4 of them full term and one at 28 weeks gestation, and found that after repeated (up to 30) "forceful" Barlow maneuvers six of the hips became lax.[85] Upon further study, it was determined that if the vacuum present in the joint capsule is disrupted, the hip becomes readily dislocatable.[85] A study of examiners with varied exam experience, using an anatomic hip model for examination, found that the maximum force applied during the Barlow maneuver far exceeded the force necessary to dislocate the joint, across all levels of experience.[86] A study with living patients used dynamic ultrasound to monitor laxity during 4 successive examinations with Barlow and Ortolani and found no increased laxity over the course of these exams.[87] However,

different examiners conducted each exam, so within-subject trends in stability were likely to reflect differences in examiners rather than changes in the joints themselves.[87]

Radiation Exposure. A single center study of radiation exposure and increased theoretical risk of fatal cancers or reproductive defects reviewed the radiographic history of 173 patients who completed a course of treatment for DDH between 1980 and 1993. Results showed that patients who had surgery (a marker for significantly more exposure) were calculated to have a 0.09% increased risk of fatal leukemia and a 0.23% increased risk of reproductive defects in males, and 0.12% and 0.5% increased risk, respectively, in females.[88] There was no increased risk of fatal breast cancer in either gender. Attributable risks in nonsurgical patients were approximately 1/2 to 1/3 of those reported for surgical patients. Given changes in technology and management in the time interval since this data was gathered, it is not clear whether the level of radiation exposure documented in this study is still applicable.

Psychosocial. We found no published studies, but identified unpublished data from Drs. Frances Gardner and Carol Dezateau on the psychosocial impacts of screening and intervention for DDH in the UK Hip Trial. This data was not made available for this review.

No evidence was identified regarding adverse effects suffered by the child or family from false positive identification. Presumably, there is a cost borne by the family and/or society for the follow-up evaluation that ensues, but this has not been quantified. Other adverse effects may be experienced, but are not represented in the literature.

Key Question 5. Does Early Diagnosis of DDH Lead to Early Intervention, and Does Early Intervention Reduce the Need for Surgery or Improve Functional Outcomes?

Family/patient adherence. Underlying the effectiveness of early diagnosis and early intervention is the degree to which families adhere to medical recommendations. One study that met quality criteria assessed failure to follow-up with a specialty appointment after identification of newborns with an abnormality on exam or the presence of a risk factor for DDH.[36] This specialty clinic, a part of Britain's National Health System, followed a systematic approach to contacting non-attenders, including up to 2 letters to the family explaining the reason for referral, safety of ultrasound, and offering an appointment the following week, followed by contact with the general practitioner to persuade the family. With this approach, nearly 95% of patients followed up. The groups with the highest follow-up rate, in excess 98%, included those with an unstable hip at the newborn exam and those with a positive family history. It may be unlikely to expect the average orthopedic clinic in the United States (US) to achieve an equivalent rate of follow-up, given established barriers to access and less robust efforts at contacting those who initially miss scheduled appointments.

A second study, based in the US, examined the rates of parental adherence to recommended abduction therapy with the Pavlik harness.[89] Of 32 patients treated by the same physician, only 2 families reported strict adherence to the physician's orders in a post-treatment questionnaire. Nonadherence was defined as failure to do one or more of the following: a) full-time use during the initial period of reduction when the hip was not stable, b) altering or deliberating misplacing the harness, c) discontinuing use of the

harness for prolonged periods of time without permission. Nearly two-thirds of the mothers in the study had a college education or advanced degree; their age range was 17-40 years (average age 29 years). Harness therapy failed in 3 out of the 32 patients, and by the authors' report these cases were not more egregious in their degree of noncompliance than successfully treated children. The single exception was a mother who routinely removed or adjusted the harness because the child could not fit into a car seat due to limited adduction.[89]

Effectiveness of interventions. A large number of nonsurgical abduction devices are represented in the published literature and an equally large number of surgical procedures are used to treat DDH (Appendix 1). The indications and timing of surgery, and the protocol for the selected treatment modality vary from site to site, further obfuscating attempts at clarifying effectiveness. These circumstances are characteristic of interventions that have not been evaluated, or proven effective, in controlled trials.[90] Because no experimental or prospective cohort studies compare intervention with no intervention, the net benefits and harms of interventions for DDH are unclear, not only for infants diagnosed early but for all children.[91]

Table 6 summarizes intervention studies[92-104] that included any assessment of functional outcomes, regardless of quality. In contrast to readily obtainable radiographic measurements of the bony anatomy of the hip joint (see below), poor functional outcomes from hip pathology may not manifest for decades. Thus, functional outcomes are not commonly measured. Even when measured, the effect of interventions on functional outcomes is unknown because of 1) the absence of an appropriate comparison cohort and 2) the substantial risk of bias stemming from short duration of follow-up, significant loss

to follow-up, and/or nonstandardized, unblinded assessment methods without adequate rigor to ensure their validity (e.g. the surgeon's subjective report of the patient's function and pain). In the absence of direct evidence from controlled trials, the case for the effectiveness of early intervention rests on less secure grounds.

Biological plausibility. It is biologically plausible that putting hips in the hip socket would facilitate normal development. While they are retrospective, careful analyses of late-presentation cases provide convincing fair quality evidence that late-presentation dislocations are often accompanied by premature arthritis, indicating that, at least in some cases, untreated DDH can have serious consequences.[105-107]

Based on this information, it is reasonable to hypothesize that relocating hips long before clinical symptoms occur may prevent morbidity and improve function. Unfortunately, an understanding of the effectiveness of interventions for DDH is confounded by the fact that many unstable and dysplastic hips undergo spontaneous resolution.[6] Thus, without a study design that includes an untreated cohort, the benefit attributable to an intervention remains in doubt.

Although the number of studies is small, it is clear that untreated DDH has an unpredictable course. Among 628 Navajo infants born in a single region from 1955 to 1961, 548 were examined and radiographed during the first four years of life (20% in the first 6 months of life, but none as neonates).[108, 109] Eighteen (3.3% of those examined) were found to have hip dysplasia (including subluxation, but not including frank dislocation) by accepted radiographic criteria. None were treated. Seventeen of these 18 children were followed for seven to 19 years, and all had stable hips with normal x-rays.[109] When 10 of these patients were followed up at 33-37 years of age, none were

aware that they had ever had a problem with their hips. While 6 did report a history of mild hip pain, this did not correlate with the degree of abnormality on x-ray. Additionally, all patients had normal function, engaged in light to heavy labor and were able to contribute to society without limitations.[108] Another study followed 51 consecutive patients with a normal clinical examination but evidence of dysplasia on x-ray. Altogether, 6 patients were lost over 5 years of follow-up. Forty-four affected hips (number of patients not reported) were normal after 5 years, 4 had undergone successful abduction therapy, and 20 were borderline on repeat imaging. No progression to subluxation or dislocation was noted in any of the hips.[110]

Reduced need for surgery. Early noninvasive intervention may reduce the need for surgery. This is a key observation that underlies several recommendations favoring screening for DDH. As discussed earlier, however (KQ1), the evidence supporting this assertion is conflicting. More over, the need for surgery is a moving target: when they are observed, reductions in surgical rates might have occurred because of changing indications or because of wider use of a period of observation prior to surgery, rather than because of screening itself.

Earlier intervention may reduce the risk of complications. In addition to studies summarized in Table 6, several observational studies examined the impact of age at the time of intervention (Key Question 5a).[32, 45, 96, 111-114] In one small study that included children initiating therapy for DDH from birth through 4 months of age, duration of treatment increased in a dose response fashion as the age at initiation of treatment increased, holding the severity of DDH steady.[45] In a separate series of patients undergoing surgery for DDH (70% of whom had failed therapy with a Pavlik harness),

those 6-9 months of age (18 patients) required no additional corrective surgeries, whereas 29% of patients 10-11 months of age, 13% of patients 12-14 months of age, 26% of patients 15-18 months of age, and 30% of patients 19-24 months of age required additional surgical interventions.[111] Another study, based upon unadjusted analysis, reported that the average age of DDH cases complicated by avascular necrosis was > 15 months, whereas uncomplicated cases averaged 11 months of age.[112] Two additional studies found that intervention initiated after 6 months of age was associated with significantly higher rates of avascular necrosis.[96, 113] In a study that focused on late diagnosis of DDH, closed reduction failed in a similar proportion of cases in children 0-3 months as those 3-6 months, but failed significantly more frequently after 6 months of age (no upper age limit could be identified in the latter category, potentially biasing these conclusions).[114] Finally, a study of 55 children who underwent operative procedures for DDH between 1988 and 1998 found that while more children diagnosed under 3 months of age underwent surgery (no denominator data was available to provide a rate), the procedures were less invasive in children less than 6 months. All children greater than 12 months undergoing a procedure for DDH required an osteotomy, the most invasive procedure.[32]

In contrast, three retrospective observational studies did not support an effect of age on success of treatment.[95, 115, 116] The first reviewed the rate of success of closed reduction, and showed no difference among patients treated with this intervention at less than 6 months, 7-12 months, or 13-18 months.[115] Next, a study limited to 168 children with hip subluxation or dislocation and a minimum follow up of 5 years, compared children in whom a Pavlik harness was successful with those requiring closed reduction

and those who eventually required open reduction, and found that age was not a predictive factor of the success of nonsurgical therapy.[116] Finally, a study of 75 children with DDH treated within the first 14 weeks of life with the Pavlik method showed that age at initiation (ranging from 5 to 13 weeks) had no influence on duration of treatment, success rate, or AVN outcome at 1 year of age.[95]

It is possible that some relevant literature was excluded because we limited the review to studies in children less than 1 year of age. However, within this age range, conclusive evidence of a clear benefit of earlier intervention is elusive. The design of the studies cannot exclude other plausible explanations for the association between age at intervention and rates of surgery. One of these explanations is that passive abduction therapy may be less effective as children become stronger and more mobile beyond 6 months of age. Another is that the early-treated group includes a high proportion of children with mild disease that would have recovered without intervention, while the older children have severe disease that would not have responded had they been treated earlier.

Improved radiographic appearance. Use of noninvasive treatments is often associated with improvements in radiographic or ultrasonographic appearance. While radiographic reduction may be an essential step in the causal pathway from congenital dislocation to prevention of serious complications, radiographic outcomes have not been shown to be valid or reliable surrogates for functional outcomes. The most commonly used and widely accepted radiographic assessment is a 6-level scale initially described by Severin in 1941, based upon radiological appearance of hips in 16-24 year olds.[117] No studies attempted to validate the Severin classification. One study examined patients

who had received surgery for dislocation of the hip, at an average of 31 years post-intervention.[118] The study found that x-ray findings (normal position of femoral neck and head, degree of arthritis and shape of the femoral head) were poorly correlated with the outcomes of range of motion and pain. Despite uncertain validity, several studies applied the Severin criteria to patients outside the range of the original 16-24 year old target population, including those not yet skeletally mature.

Two studies assessed the reliability of the Severin classification.[119, 120] Ali et al found intraobserver reliability among pediatric orthopedists in the UK with 7 or more years experience to be moderate to substantial (kappa ranging from .58 to .77), and interobserver reliability to be poor to slight in the intermediate Severin classes of II and III (kappa 0.19 to 0.20) and moderate (kappa 0.44 to 0.54) in the disparate Severin classifications of I (normal) and V (marginal dislocation). Unfortunately, "good" outcomes are typically classified as Severin II,[91] one of the grades found to have the poorest inter-observer reliability. A study by Ward found even less reassuring results.[120] Blinded assessments by pediatric orthopedists in this study were assessed by dichotomous observer groups as well as multi-rater groups, and found kappa scores in the range of 0.0 to 0.29 across the range of Severin classes, and no higher than 0.56 for overall agreement across any two surgeons. Even more concerning, the operating surgeons' unblended scores showed uniform poor reliability (kappa 0.02 to 0.21) when compared to each of the blinded observer's scores. Despite uncertain reliability, intervention studies rarely included blinded or repeated assessments of radiographic outcomes. Due to highly suspect validity and reliability, studies that reported only radiographic outcomes were excluded from further review.

Closer follow-up. Diagnosis leads to attentive follow-up of infants with DDH, facilitating quick detection and intervention. Thus, children undergoing early noninvasive therapy may benefit from closer follow-up and the physician's ability to react to a deteriorating condition more rapidly. As discussed above, available evidence supports the notion that a high proportion of families follow through with initial referral. However, we could not determine how many families adhere to ongoing follow-up.

Key Question 6. What Are the Adverse Effects of Early Diagnosis and/or Intervention?

Good quality literature examining harms of intervention for DDH would include a comparison of 2 or more (ideally randomized) cohorts, each exposed to a standardized intervention and followed over sufficient time (with limited loss to follow-up) to ensure complete ascertainment of the potential harms with an assessment of the effect of the measured harms on patient outcomes. Unfortunately, these studies have not yet been conducted. In their absence, we reviewed the fair quality literature on adverse effects of both nonsurgical and surgical interventions.

The most well described adverse effect from interventions aimed at treating DDH is AVN of the femoral head. This is the most common adverse effect for both abduction therapy and surgical interventions. AVN severity ranges from a persistent but asymptomatic radiographic finding to a severe condition that causes growth arrest and can lead to eventual destruction of the joint. The rates described in the literature for this adverse effect vary greatly for abduction therapy as well as surgical interventions. (Figure 5).[92-96, 98, 100, 102-104, 113, 121-129] The reasons for these disparate findings are not

straightforward, and most likely relate to a complex and confounded set of variables including but not limited to the wide spectrum of the disorder, heterogeneous populations studied (age at intervention, specific type of DDH, previous interventions received), the variety of interventions and the poorly standardized approach to interventions (particularly the pre- and post-intervention phase of management), variable training and talent among the treating physicians, different lengths of follow-up across studies, and disparate approaches to follow-up in different health care systems. As calculated in the AAP review, meta-analytic rates of AVN range from 13.5 - 109/ 1000 infants who undergo treatment (non-surgical vs. surgical rates not specified).[2]

Additional harms from abduction therapy that have been addressed in the literature are typically mild and self-limited, and include rash, pressure sores, and femoral nerve palsy**.** All surgical interventions carry the risks inherent in general anesthesia, and those that involve open surgery also include the generic surgical risks of infection, excessive bleeding, and wrong site surgery, though these receive scant review in the published literature and thus cannot be quantified.

A fair quality study assessing the long-term psychological impact on children of successfully treated DDH showed that parents and teachers found that children with DDH were more "disordered" than peers with no hospitalizations, 1 hospitalization, and multiple hospitalizations on measures of *health*, *habits*, and *behavior*.[130] This study which took place in 1983, implies (but does not quantify) extended hospitalizations for these children as a rule, and thus may not be generalizable to the impact of treatment today.

Key Question 7. What Cost-Effectiveness Issues Apply to Screening for DDH?

Several economic analyses of screening for DDH have been published.[48, 91, 131-136] Most concern the marginal benefit of ultrasound screening in relation to screening with clinical examination.[48, 91, 132, 133, 136] None of the available studies used quality adjusted life years, and none used models based upon U.S. data or the U.S. health care system. These analyses demonstrate that the economic impact of ultrasound screening is complex, reflecting that ultrasound may have mixed effects on diagnosis of DDH: it may identify false positive clinical examinations, reducing or shortening the duration of unnecessary treatments, but it also identifies many abnormalities in infants who have normal physical examinations, potentially leading to more early treatment and greater follow-up costs. The mixed results of the economic studies largely reflect mixed results of the clinical studies on which they are based. The best quality economic study, derived from a RCT (in the UK) of clinical exam screening versus clinical exam plus ultrasound, maintained detailed records of utilization of medical services and related costs.[48] The authors concluded that the overall direct medical costs for the two approaches were not statistically significantly different.[48] This study did not report indirect costs, such as missed work by the family, nor did it include the costs of long-term follow-up or complications.

Chapter 4. Discussion

Conclusions and Limitations of the Literature

As a condition that can result in impaired functional outcomes for children and adults, DDH merits the attention of primary care clinicians. However, as shown in Table 7, there is no direct evidence that screening improves functional outcomes, and the evidence for several links in the analytic framework is weak.

The definition of DDH is variable, including dislocated, dislocatable, subluxable, and dysplastic hips. The benefits of early intervention are based on expert opinion along with fair evidence that later diagnosis results in worse outcomes and greater need for surgical intervention. Using indirect comparisons, some studies suggest that earlier diagnosis is associated with better outcomes, but these findings could be the result of lead-time bias, that is, the identification of DDH in a group of younger patients, in whom a higher rate of spontaneous resolution may lead to better outcomes, versus the effect of earlier intervention. The outcomes of screened infants have not been compared to those of unscreened infants in an experimental or observational study.

Despite a paucity of direct evidence supporting its value in improving outcomes, universal screening for DDH is a well-established approach to the disorder. However, the approach to screening varies significantly. In addition to physical examination with the provocative tests of Barlow and Ortolani and evaluation of range of motion emphasizing abduction of the hip, static and dynamic ultrasound are employed to identify anatomic abnormalities and stability of the hip, respectively.

Some have recommended risk stratification to inform selective use of ultrasound, with females in breech positioning at delivery found to have the highest rate of clinical

hip instability (84/1000) and subsequent rate of treatment (20/1000). Other health systems have elected to employ universal ultrasound screening in an effort to reduce the incidence of late diagnosis of DDH (variably defined in the literature as diagnosis beyond 1, 3, 6, or 12 months of age). The use of ultrasound to further evaluate hips found to be unstable on clinical exam may reduce the rate of unnecessary treatment. However, the reliability of the approach to DDH classification by ultrasound is questionable. Theoretical harms from screening include examiner induced hip pathology with vigorous provocative testing, elevated risk of certain cancers from increased radiation exposure from follow-up radiographic tests, and parental psychosocial stress from the diagnosis and therapy. None of these has been quantified in patients/families in clinical studies published to date beyond anecdotes.

It is known that a significant number of hips with positive screening tests, both by physical examination and by ultrasound, will normalize over time without intervention. This is particularly true of ultrasound in hips that are stable on clinical exam of the neonate: more than 90% of abnormal ultrasound findings in this situation have been shown to normalize spontaneously. While limited fair quality evidence exists to support the value of initiating treatment within the first 6 months of life, there is little to suggest that immediate treatment in the newborn period is associated with improved outcomes or a reduced need for subsequent surgery. However, no study has examined the effect of timing of treatment initiation, controlling for the degree of instability (dysplastic, subluxated, dislocatable, dislocated).

First-line intervention includes abduction bracing of the hips, which attempts to induce passive alignment of the hip. Several devices are used for abduction, with a wide

range of institutional protocols. Failure of abduction therapy, or the occasional case of dislocated and clinically irreducible hips at presentation, leads to surgical intervention, including one or more of the following: closed reduction under general anesthesia, open reduction, and femoral and/or acetabular osteotomies. The indications and protocols for surgery vary widely, as do the pre- and post-operative approaches to management, which include traction, hip spica casting, and/or various forms of abduction bracing for extended periods.

Estimates of the effectiveness of therapy are confounded by spontaneous resolution of hip dysplasia, which has only rarely been assessed and never in a prospective or comparative fashion. The impact of interventions on functional outcomes is rarely addressed in the literature, and when addressed is of poor quality due to a lack of standardization within studies, and the absence of validated functional outcome measures across studies.

The most significant and common adverse effect of both nonsurgical and surgical intervention for hip dysplasia is avascular necrosis of the femoral head, which can lead to growth arrest and eventual destruction of the hip joint. Assessment of the cost effectiveness of screening for DDH requires more conclusive information about effectiveness. Studies including cost data on various screening programs have been conducted outside the U.S. and may not be generalizable to health care delivery in this country.

Future Research

While the body of literature on screening and intervention for DDH has significant flaws, several recent studies provide valuable information on the screening evaluation of DDH. A more complete understanding of the natural history of spontaneous resolution of hip instability and dysplasia is needed to develop an evidence-based strategy for conducting screening and implementing therapy at the optimal time. Given the infrequent nature of DDH, multicenter studies of interventions that measure functional outcomes in a standardized fashion are needed. Studies designed to assess whether any clearly defined, reliable radiological markers predict functional outcomes would be a valuable step. Even more valuable would be patient-centered research that seeks to understand patient and family preferences as they relate to the process of care and short and long-term outcomes of DDH.

References

1. Patel H, Canadian Task Force on Preventive Health C. Preventive health care, 2001 update: screening and management of developmental dysplasia of the hip in newborns. Can Med Assoc Jl. 2001;164(12):1669-1677.
2. Lehmann HP, Hinton R, Morello P, et al. . Developmental dysplasia of the hip practice guideline: technical report. Committee on Quality Improvement, and Subcommittee on Developmental Dysplasia of the Hip. Pediatrics. 2000;105(4):E57.
3. Clinical practice guideline: early detection of developmental dysplasia of the hip. Committee on Quality Improvement, Subcommittee on Developmental Dysplasia of the Hip. American Academy of Pediatrics. Pediatrics. 2000;105(4 Pt 1):896-905.
4. Bialik V, Bialik GM, Blazer S, et al. Developmental dysplasia of the hip: a new approach to incidence. Pediatrics. 1999;103(1):93-99.
5. Gardiner HM, Clarke NM, Dunn PM. A sonographic study of the morphology of the preterm neonatal hip. J of Pediatr Ortho. 1990;10(5):633-637.
6. Barlow T. Early diagnosis and treatment of congenital dislocation of the hip. J Bone Joint Surg. 1962;44:292-301.
7. Omeroglu H, Koparal S. The role of clinical examination and risk factors in the diagnosis of developmental dysplasia of the hip: a prospective study in 188 referred young infants. Arch Ortho Trauma Surg. 2001;121(1-2):7-11.
8. Yiv BC, Saidin R, Cundy PJ, et al. Developmental dysplasia of the hip in South Australia in 1991: prevalence and risk factors. J Paediatr Child Health. 1997;33(2):151-156.
9. Chan A, McCaul KA, Cundy PJ, et al. Perinatal risk factors for developmental dysplasia of the hip. Arch Dis Child. 1997;76(2):F94-100.
10. Standing Medical Advisory Committee, Standing Nursing and Midwifery Advisory Committee Working Party for the Secretaries of State for Social Services and Wales. Screening for the detection of congenital dislocation of the hip. Arch Dis Child. 1986;61(9):921-926.
11. Barlow T. Early diagnosis and treatment of congenital dislocation of the hip in the newborn. Proc R SocMed. 1966;59:1103-1106.

12. Ortolani M. Congenital hip dysplasia in the light of early and very early diagnosis. Clin Orthop Relat Res. 1976(119):6-10.

13. Baronciani D, Atti G, Andiloro F, et al. Screening for developmental dysplasia of the hip: from theory to practice. Collaborative Group DDH Project. Pediatrics.1997;99(2):E5.

14. Bialik V, Bialik GM, Wiener F. Prevention of overtreatment of neonatal hip dysplasia by the use of ultrasonography. J Pediatr Orthop, Part B. 1998;7(1):39-42.

15. Andersson JE. Neonatal hip instability: results and experiences from ten years of screening with the anterior-dynamic ultrasound method. Acta Paediatrica. 2002;91(8):926-929.

16. Giannakopoulou C, Aligizakis A, Korakaki E, et al. Neonatal screening for developmental dysplasia of the hip on the maternity wards in Crete, Greece. Correlation to risk factors. Clin Exper ObstGyn. 2002;29(2):148-152.

17. Vitale MG, Skaggs DL. Developmental dysplasia of the hip from six months to four years of age. J Am Acad Orthop Surg. 2001;9(6):401-411.

18. Slavin R. Best practice synthesis: an alternative to meta-analytic and traditional reviews. Educ Res. 1986;15:5-11.

19. Harris RP, Helfand M, Woolf SH, et al. Current methods of the third U.S. Preventive Services Task Force. Am J Prev Med. 2001;20(3S):21-35.

20. McDonagh M, Carson S, Ash J, al. e. Hyperbaric oxygen therapy for brain injury, cerebral palsy, and stroke. Evidence Report/Technology Assessment No. 85 Rockville, MD: Agency for Healthcare, Research, and Quality; 2003 September. AHRQ Publication No. 04-E003.

21. Rosendahl K, Markestad T, Lie RT. Ultrasound screening for developmental dysplasia of the hip in the neonate: the effect on treatment rate and prevalence of late cases. Pediatrics. 1994;94(1):47-52.

22. Tredwell SJ, Bell HM. Efficacy of neonatal hip examination. JPediatr Ortho. 1981;1(1):61-65.

23. Dunn PM, Evans RE, Thearle MJ, et al. Congenital dislocation of the hip: early and late diagnosis and management compared. ArchDisChild. 1985;60(5):407-414.

24. Yngve D, Gross R. Late diagnosis of hip dislocation in infants. J Pediatr Ortho. 1990;10(6):777-779.

25. Macnicol MF. Results of a 25-year screening programme for neonatal hip instability. J Bone Joint Surg - British Volume. 1990;72(6):1057-1060.

26. Place MJ, Parkin DM, Fritton JMKS-. Effectiveness of neonatal screening for congenital dislocation of the hip. Lancet. 1978;2(8083):249-250.

27. MacKenzie IG, Wilson JG. Problems encountered in the early diagnosis and management of congenital dislocation of the hip. J Bone Joint Surg - British Volume. 1981;63-B(1):38-42.

28. Sanfridson J, Redlund-Johnell I, Uden A. Why is congenital dislocation of the hip still missed? Analysis of 96,891 infants screened in Malmo 1956-1987. Acta Orthopaedica Scandinavica. 1991;62(2):87-91.

29. Catford JC, Bennet GC, Wilkinson JA. Congenital hip dislocation: an increasing and still uncontrolled disability? Br Med J (Clin Res Ed). 1982;285(6354):1527-1530.

30. Patterson CC, Kernohan WG, Mollan RA, Haugh PE, Trainor BP. High incidence of congenital dislocation of the hip in Northern Ireland. Paediatr Perinat Epidemiol. 1995;9(1):90-97.

31. Godward S, Dezateux C. Surgery for congenital dislocation of the hip in the UK as a measure of outcome of screening. MRC Working Party on Congenital Dislocation of the Hip. Medical Research Council. Lancet. 1998;351(9110):1149-1152.

32. Chan A, Cundy PJ, Foster BK, et al. Late diagnosis of congenital dislocation of the hip and presence of a screening programme: South Australian population-based study. Lancet. 1999;354(9189):1514-1517.

33. Andersson JE, Oden A. The breech presentation and the vertex presentation following an external version represent risk factors for neonatal hip instability. Acta Paediatr. 2001;90:895-898.

34. Artz TD, Lim WN, Wilson PD, Levine DB, Salvati EA. Neonatal diagnosis, treatment and related factors of congenital dislocation of the hip. Clin Orthop Relat Res. 1975(110):112-136.

35. Bache CE, Clegg J, Herron M. Risk factors for developmental dysplasia of the hip: ultrasonographic findings in the neonatal period. J Pediatr Ortho, Part B. 2002;11(3):212-218.

36. Boeree NR, Clarke NM. Ultrasound imaging and secondary screening for congenital dislocation of the hip. J Bone Joint Surg - British Volume. 1994;76(4):525-533.

37. Boere-Boonekamp MM, Kerkhoff TH, Schuil PB, . Early detection of developmental dysplasia of the hip in The Netherlands: the validity of a standardized assessment protocol in infants. Am J Public Health. 1998;88(2):285-288.

38. Goss PW. Successful screening for neonatal hip instability in Australia. J PaediatrChild Health. 2002;38(5):469-474.

39. Holen KJ, Tegnander A, Terjesen T, et al. Ultrasonographic evaluation of breech presentation as a risk factor for hip dysplasia. Acta Paediatr. 1996;85(2):225-229.

40. Jones DA. Importance of the clicking hip in screening for congenital dislocation of the hip. Lancet. 1989;1(8638):599-601.

41. Miranda L, Palomo JM, Monzonis J, et al. Prevention of congenital dislocation of the hip in the newborn. J Pediatr Ortho. 1988;8(6):671-675.

42. Sahin F, Akturk A, Beyazova U, et al. Screening for developmental dysplasia of the hip: results of a 7-year follow-up study. Pediatr Int. 2004;46(2):162-166.

43. Walter RS, Donaldson JS, Davis CL, et al. Ultrasound screening of high-risk infants. A method to increase early detection of congenital dysplasia of the hip. Am J Dis Child. 1992;146(2):230-234.

44. Paton RW, Srinivasan MS, Shah B, et al. Ultrasound screening for hips at risk in developmental dysplasia. Is it worth it? J Bone Joint Surg - British Volume. 1999;81(2):255-258.

45. Riboni G, Bellini A, Serantoni S, et al. Ultrasound screening for developmental dysplasia of the hip. Pediatr Radiol. 2003;33(7):475-481.

46. Rosenberg N, Bialik V, Norman D, et al. The importance of combined clinical and sonographic examination of instability of the neonatal hip. Int Ortho. 1998;22(3):185-188.

47. Rosendahl K, Markestad T, Lie RT. Developmental dysplasia of the hip. A population-based comparison of ultrasound and clinical findings. Acta Paediatr. 1996;85(1):64-69.

48. Elbourne D, Dezateux C, Arthur R, et al. Ultrasonography in the diagnosis and management of developmental hip dysplasia (UK Hip Trial): clinical and economic results of a multicentre randomised controlled trial. Lancet. 2002;360(9350):2009-2017.

49. Holen KJ, Tegnander A, Bredland T, et al. Universal or selective screening of the neonatal hip using ultrasound? A prospective, randomised trial of 15,529 newborn infants. J Bone Joint Surg - British Volume. 2002;84(6):886-890.

50. Bjerkreim I, Arseth PH. Congenital dislocation of the hip in Norway. Late diagnosis CDH in the years 1970 to 1974. Acta Paediatr Scan. 1978;67(3):329-332.

51. Quinland WR, Brady PG, Regan BF. Late diagnosis of congenital dislocation of the hip. Ir Med J. 1978;71(13):447-449.

52. Glass A, Dunningham TH. Capsular arthroplasty of the hip for congenital dislocation. Isr J Med Sci. 1980;16(4):328-332.

53. Kepley RF, Weiner DS. Treatment of congenital dysplasia-subluxation of the hip in children under one year of age. J Pediatr Ortho. 1981;1(4):413-418.

54. Late diagnosis of hip dislocation. Lancet. 1983;2(8342):145.

55. Heikkila E, Ryoppy S, Louhimo I. Late diagnosis in congenital dislocation of the hip. Acta Ortho Scan. 1984;55(3):256-260.

56. Lehmann EC, Street DG. Neonatal screening in Vancouver for congenital dislocation of the hip. Can Med Assoc J. 1981;124(8):1003-1008.

57. Watanabe M, Yanagisawa M, Fukuda S, Takahashi I, Matsumoto M. [Examination, prevention and treatment of congenital dislocation of the hip in the newborn infant--experience over an eighteen-year period]. Nippon Seikeigeka Gakkai Zasshi - J Jap Ortho Assoc. 1986;60(11):1063-1078.

58. Knox EG, Armstrong EH, Lancashire RJ. Effectiveness of screening for congenital dislocation of the hip. J Epidemiol Community Health. 1987;41(4):283-289.

59. Watanabe M, Yanagisawa M. Late diagnosis of congenital dislocation of the hip in the newborn. Fukushima J Med Sci. 1988;34(2):75-80.

60. Hazel JR, Beals RK. Diagnosing dislocation of the hip in infancy. West J Med. 1989;151(1):39-41.

61. Kernohan WG, Nugent GE, Haugh PE, et al. Sensitivity of manual palpation in testing the neonatal hip. Clin Orthop Relat Res. 1993(294):211-215.

62. Clarke NM, Clegg J, Al-Chalabi AN. Ultrasound screening of hips at risk for CDH. Failure to reduce the incidence of late cases. J Bone Joint Surg - British Volume. 1989;71(1):9-12.

63. Sampath JS, Deakin S, Paton RW. Splintage in developmental dysplasia of the hip: how low can we go? J Pediatr Ortho. 2003;23(3):352-355.

64. Castelein RM, Sauter AJ, de Vlieger M, et al. Natural history of ultrasound hip abnormalities in clinically normal newborns. J Pediatr Ortho. 1992;12(4):423-427.

65. Terjesen T, Holen KJ, Tegnander A. Hip abnormalities detected by ultrasound in clinically normal newborn infants. J Bone Joint Surg - British Volume. 1996;78(4):636-640.

66. Wood MK, Conboy V, Benson MK. Does early treatment by abduction splintage improve the development of dysplastic but stable neonatal hips? J Pediatr Ortho. 2000;20(3):302-305.

67. Burger BJ, Burger JD, Bos CF, et al. Neonatal screening and staggered early treatment for congenital dislocation or dysplasia of the hip. Lancet. 1990;336(8730):1549-1553.

68. Marks DS, Clegg J, al-Chalabi AN. Routine ultrasound screening for neonatal hip instability. Can it abolish late-presenting congenital dislocation of the hip? J Bone Joint Surg - British Volume. 1994;76(4):534-538.

69. Puhan MA, Woolacott N, Kleijnen J, et al. Observational studies on ultrasound screening for developmental dysplasia of the hip in newborns - a systematic review. Ultraschall in der Medizin. 2003;24(6):377-382.

70. Palmen K. Preluxation of the hip joint. Diagnosis and treatment in the newborn and the diagnosis of congenital dislocation of the hip joint in Sweden during the years 1948-1960. Acta Paediatr. 1961;50(Suppl 129):1-71.

71. Palmen K. Prevention of congenital dislocation of the hip. The Swedish experience of neonatal treatment of hip joint instability. Acta Ortho Scan. Supplementum. 1984;208:1-107.

72. Graf R. Fundamentals of sonographic diagnosis of infant hip dysplasia. J Pediatr Ortho. 1984;4(6):735-740.

73. Bar-On E, Meyer S, Harari G, et al. Ultrasonography of the hip in developmental hip dysplasia. J Bone Joint Surg - British Volume. 1998;80(2):321-324.

74. Rosendahl K, Aslaksen A, Lie RT, et al. Reliability of ultrasound in the early diagnosis of developmental dysplasia of the hip. Pediatr Radiol. 1995;25(3):219-224.

75. Dias JJ, Thomas IH, Lamont AC, et al. The reliability of ultrasonographic assessment of neonatal hips. J Bone Joint Surg - British Volume. 1993;75(3):479-482.

76. Jari S, Paton RW, Srinivasan MS. Unilateral limitation of abduction of the hip. A valuable clinical sign for DDH? J Bone Joint Surg - British Volume. 2002;84(1):104-107.

77. Castelein RM, Korte J. Limited hip abduction in the infant. J Pediatr Ortho. 2001;21(5):668-670.

78. Hansson G, Romanus B, Scheller S. Pitfalls of early diagnosis and treatment of congenital dislocation of the hip joint. Arch Ortho Traumatic Surg. 1988;107(3):129-135.

79. Bialik V, Fishman J, Katzir J, et al. Clinical assessment of hip instability in the newborn by an orthopedic surgeon and a pediatrician. J Pediatr Ortho. 1986;6(6):703-705.

80. Holen KJ, Tegnander A, Terjesen T, et al. Ultrasonography of clinically unstable hips. A prospective study of 143 neonates at birth and early follow-up. Acta Ortho Scan. 1997;68(6):527-532.

81. Lee TW, Skelton RE, Skene C. Routine neonatal examination: effectiveness of trainee paediatrician compared with advanced neonatal nurse practitioner. Arch Dis Child. 2001;85(2):F100-104.

82. Fiddian NJ, Gardiner JC. Screening for congenital dislocation of the hip by physiotherapists. Results of a ten-year study. J Bone Joint Surg - British Volume. 1994;76(3):458-459.

83. Krikler SJ, Dwyer NS. Comparison of results of two approaches to hip screening in infants. J Bone Joint Surg - British Volume. 1992;74(5):701-703.

84. Moore FH. Examining infants' hips--can it do harm? J Bone Joint Surg - British Volume. 1989;71(1):4-5.

85. Jones DA. Neonatal hip stability and the Barlow test. A study in stillborn babies. J Bone Joint Surg - British Volume. 1991;73(2):216-218.

86. Chow YW, Turner I, Kernohan WG, et al. Measurement of the forces and movements involved in neonatal hip testing. Med Eng Phys. 1994;16(3):181-187.

87. Andersson JE. Neonatal hip instability: normal values for physiological movement of the femoral head determined by an anterior-dynamic ultrasound method. J Pediatr Ortho. 1995;15(6):736-740.

88. Bone CM, Hsieh GH. The risk of carcinogenesis from radiographs to pediatric orthopaedic patients. J Pediatr Ortho. 2000;20(2):251-254.

89. McHale KA, Corbett D. Parental noncompliance with Pavlik harness treatment of infantile hip problems. J Pediatr Ortho. 1989;9(6):649-652.

90. Bunker JP, Barnes BA, Mosteller F, editors. Costs, risks, and benefits of surgery. New York: Oxford University Press; 1977.

91. Dezateux C, Brown J, Arthur R, et al. Performance, treatment pathways, and effects of alternative policy options for screening for developmental dysplasia of the hip in the United Kingdom. Arch Dis Child. 2003;88(9):753-759.

92. Aksoy MC, Ozkoc G, Alanay A, et al. Treatment of developmental dysplasia of the hip before walking: results of closed reduction and immobilization in hip spica cast. Turk J Pediatr. 2002;44(2):122-127.

93. Cashman JP, Round J, Taylor G, et al. The natural history of developmental dysplasia of the hip after early supervised treatment in the Pavlik harness. A prospective, longitudinal follow-up. J Bone Joint Surg - British Volume. 2002;84(3):418-425.

94. Danielsson L. Late-diagnosed DDH: a prospective 11-year follow-up of 71 consecutive patients (75 hips). Acta Ortho Scan. 2000;71(3):232-242.

95. Eidelman M, Katzman A, Freiman S, et al. Treatment of true developmental dysplasia of the hip using Pavlik's method. J Pediatr Ortho. 2003;12(4):253-258.

96. Konigsberg DE, Karol LA, Colby S, et al. Results of medial open reduction of the hip in infants with developmental dislocation of the hip. J Pediatr Ortho. 2003;23(1):1-9.

97. Gregersen HN. Congenital dislocation of the hip. Results of early treatment. Acta Ortho Scan. 1969;40(1):53-61.

98. Malvitz TA, Weinstein SL. Closed reduction for congenital dysplasia of the hip. Functional and radiographic results after an average of thirty years. J Bone Joint Surg - American Volume. 1994;76(12):1777-1792.

99. O'Hara JN, Bernard AA, Dwyer NS. Early results of medial approach open reduction in congenital dislocation of the hip: use before walking age. J Pediatr Ortho. 1988;8(3):288-294.

100. Sosna A, Rejholec M. Ludloff's open reduction of the hip: long-term results. J Pediatr Ortho. 1992;12(5):603-606.

101. Tegnander A, Holen KJ, Anda S, et al. Good results after treatment with the Frejka pillow for hip dysplasia in newborns: a 3-year to 6-year follow-up study. J Pediatr Ortho. 2001;10(3):173-179.

102. Tumer Y, Ward WT, Grudziak J. Medial open reduction in the treatment of developmental dislocation of the hip. J Pediatr Ortho. 1997;17(2):176-180.

103. Yamada N, Maeda S, Fujii G, et al. Closed reduction of developmental dislocation of the hip by prolonged traction. J Bone Joint Surg - British Volume. 2003;85(8):1173-1177.

104. Yoshitaka T, Mitani S, Aoki K, et al. Long-term follow-up of congenital subluxation of the hip. J Pediatr Ortho. 2001;21(4):474-480.

105. Wedge JH, Wasylenko MJ. The natural history of congenital disease of the hip. J Bone Joint Surg - British Volume. 1979;61-B(3):334-338.

106. Wenger DR, Frick SL. Early surgical correction of residual hip dysplasia: the San Diego Children's Hospital approach. Acta Ortho Belgica. 1999;65(3):277-287.

107. Cooperman DR, Wallensten R, Stulberg SD. Post-reduction avascular necrosis in congenital dislocation of the hip. J Bone Joint Surg - American Volume. 1980;62(2):247-258.

108. Schwend RM, Pratt WB, Fultz J. Untreated acetabular dysplasia of the hip in the Navajo. A 34 year case series followup. Clin Orthop Relat Res. 1999(364):108-116.

109. Pratt WB, Freiberger RH, Arnold WD. Untreated congenital hip dysplasia in the Navajo. Clin Orthop Relat Res. 1982(162):69-77.

110. Mladenov K, Dora C, Wicart P, et al. Natural history of hips with borderline acetabular index and acetabular dysplasia in infants. J Pediatr Ortho. 2002;22(5):607-612.

111. Szepesi K, Biro B, Fazekas K, et al. Preliminary results of early open reduction by an anterior approach for congenital dislocation of the hip. J Pediatr Ortho. 1995;4(2):171-178.

112. Castillo R, Sherman FC. Medial adductor open reduction for congenital dislocation of the hip. J Pediatr Ortho. 1990;10(3):335-340.

113. Kruczynski J. Avascular necrosis of the proximal femur in developmental dislocation of the hip. Incidence, risk factors, sequelae and MR imaging for diagnosis and prognosis. Acta Ortho Scan. Supp. 1996;268:1-48.

114. Maxwell SL, Ruiz AL, Lappin KJ, et al. Clinical screening for developmental dysplasia of the hip in Northern Ireland. BMJ. 2002;324(7344):1031-1033.

115. Bicimoglu A, Agus H, Omeroglu H, et al. Six years of experience with a new surgical algorithm in developmental dysplasia of the hip in children under 18 months of age. J Pediatr Ortho. 2003;23(6):693-698.

116. Takahashi I. Functional treatment of congenital dislocation of the hip using Pavlik harness (Riemenbugel). Nippon Seikeigeka Gakkai Zasshi - J Jap Ortho Assoc. 1985;59(11):973-984.

117. Severin E. Contribution to the knowledge of congenital dislocation of the hip: late results of closed reduction and arthrographic studies of recent cases. Acta Chir Scand. 1941;84 Supplementum 63:1-142.

118. Smith WS, Badgley CE, Orwig JB, et al. Correlation of postreduction roentgenograms and thirty-one-year follow-up in congenital dislocation of the hip. J Bone Joint Surg - American Volume. 1968;50(6):1081-1098.

119. Ali AM, Angliss R, Fujii G, et al. Reliability of the Severin classification in the assessment of developmental dysplasia of the hip. J Pediatr Ortho. 2001;10(4):293-297.

120. Ward WT, Vogt M, Grudziak JS, et al. Severin classification system for evaluation of the results of operative treatment of congenital dislocation of the hip. A study of intraobserver and interobserver reliability. J Bone Joint Surg - American Volume. 1997;79(5):656-663.

121. Brougham DI, Broughton NS, Cole WG, et al. Avascular necrosis following closed reduction of congenital dislocation of the hip. Review of influencing factors and long-term follow-up. J Bone Joint Surg - British Volume. 1990;72(4):557-562.

122. Buchanan JR, Greer RB, 3rd, Cotler JM. Management strategy for prevention of avascular necrosis during treatment of congenital dislocation of the hip. J Bone Joint Surg - American Volume. 1981;63(1):140-146.

123. Grill F, Bensahel H, Canadell J, et al. The Pavlik harness in the treatment of congenital dislocating hip: report on a multicenter study of the European Paediatric Orthopaedic Society. J Pediatr Ortho.1988;8(1):1-8.

124. Lennox IA, McLauchlan J, Murali R. Failures of screening and management of congenital dislocation of the hip. J Bone Joint Surg - British Volume. 1993;75(1):72-75.

125. Pool RD, Foster BK, Paterson DC. Avascular necrosis in congenital hip dislocation. The significance of splintage. J Bone Joint Surg - British Volume. 1986;68(3):427-430.

126. Powell EN, Gerratana FJ, Gage JR. Open reduction for congenital hip dislocation: the risk of avascular necrosis with three different approaches. J Pediatr Ortho. 1986;6(2):127-132.

127. Suzuki S, Seto Y, Futami T, et al. Preliminary traction and the use of under-thigh pillows to prevent avascular necrosis of the femoral head in Pavlik harness treatment of developmental dysplasia of the hip. J Orthop Sci. 2000;5(6):540-545.

128. Thomas IH, Dunin AJ, Cole WG, et al. Avascular necrosis after open reduction for congenital dislocation of the hip: analysis of causative factors and natural history. J Pediatr Ortho. 1989;9(5):525-531.

129. Weiner DS. Avascular necrosis as a treatment complication in congenital dislocation of the hip in children under one year of age. Isr J Med Sci. 1980;16(4):301-306.

130. Standen PJ. The long-term psychological adjustment of children treated for congenital dislocation of the hip. Psychol Med. 1983;13(4):847-854.

131. Brown J, Dezateux C, Karnon J, et al. Efficiency of alternative policy options for screening for developmental dysplasia of the hip in the United Kingdom. Arch Dis Child. 2003;88(9):760-766.

132. Bralic I, Vrdoljak J, Kovacic L. Ultrasound screening of the neonatal hip: cost-benefit analysis. Croat Med J. 2001;42(2):171-174.

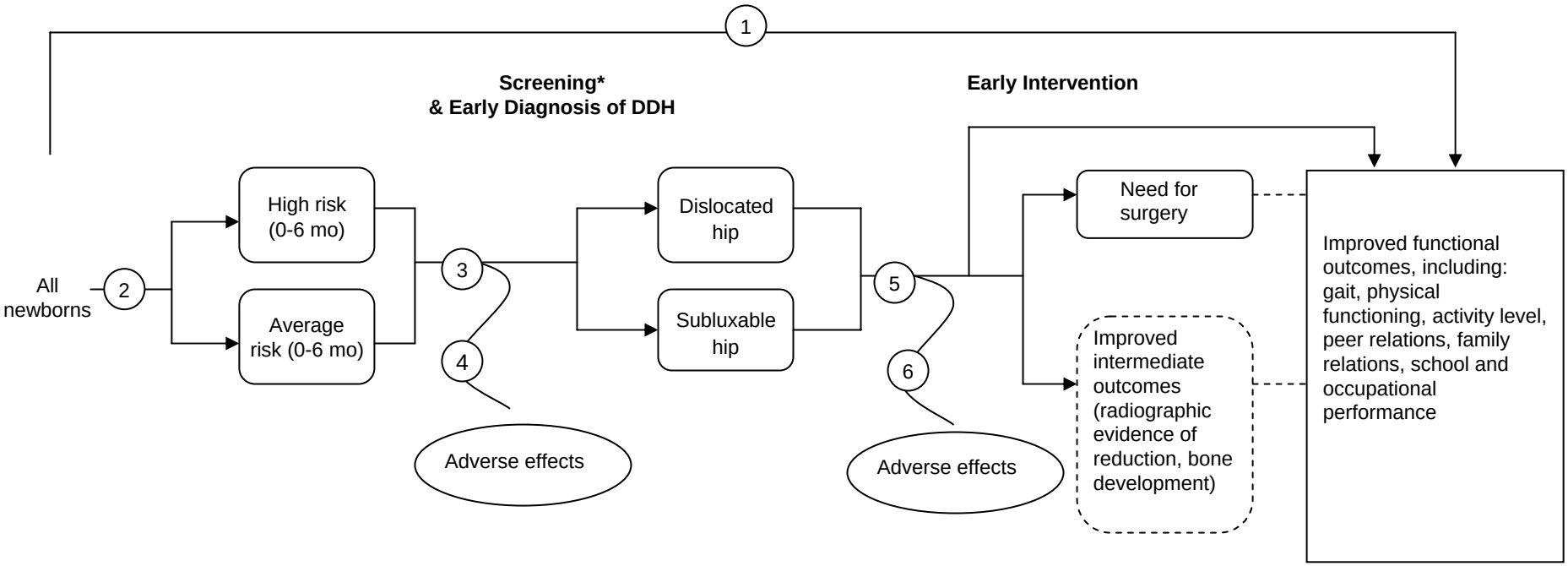
133. Clegg J, Bache CE, Raut VV. Financial justification for routine ultrasound screening of the neonatal hip. J Bone Joint Surg - British Volume. 1999;81(5):852-857.

134. Faure C, Schmit P, Salvat D. Cost-benefit evaluation of systematic radiological diagnosis of congenital dislocated hip. Pediatr Radiol. 1984;14(6):407-412.

135. Rosendahl K, Markestad T, Lie RT, et al. Cost-effectiveness of alternative screening strategies for developmental dysplasia of the hip. Arch PediatrAdolesc Med. 1995;149(6):643-648.

136. Geitung JT, Rosendahl K, Sudmann E. Cost-effectiveness of ultrasonographic screening for congenital hip dysplasia in new-borns. Skeletal Radiol. 1996;25(3):251-254.

Figure 1. Analytic Framework



*Screening examination (Barlow/Ortolani, Asymmetry, ROM) and Radiographic evaluation (ultrasound, x-ray)

Figure 2. Key Questions

KQ 1: Does screening for DDH lead to improved outcomes (including reduced need for surgery and improved functional outcomes such as: gait, physical functioning, activity level, peer relations, family relations, school and occupational performance)?

KQ 2: Can infants at high risk for DDH be identified, and does this group warrant a different approach to screening than children at average risk?

KQ 3: Does screening for DDH lead to early identification of children with DDH?

a) What is the accuracy of clinical examination and ultrasound?

b) How does the age of the child affect screening parameters?

c) How does the educational level and training of the screener impact screening?

KQ 4: What are the adverse effects of screening?

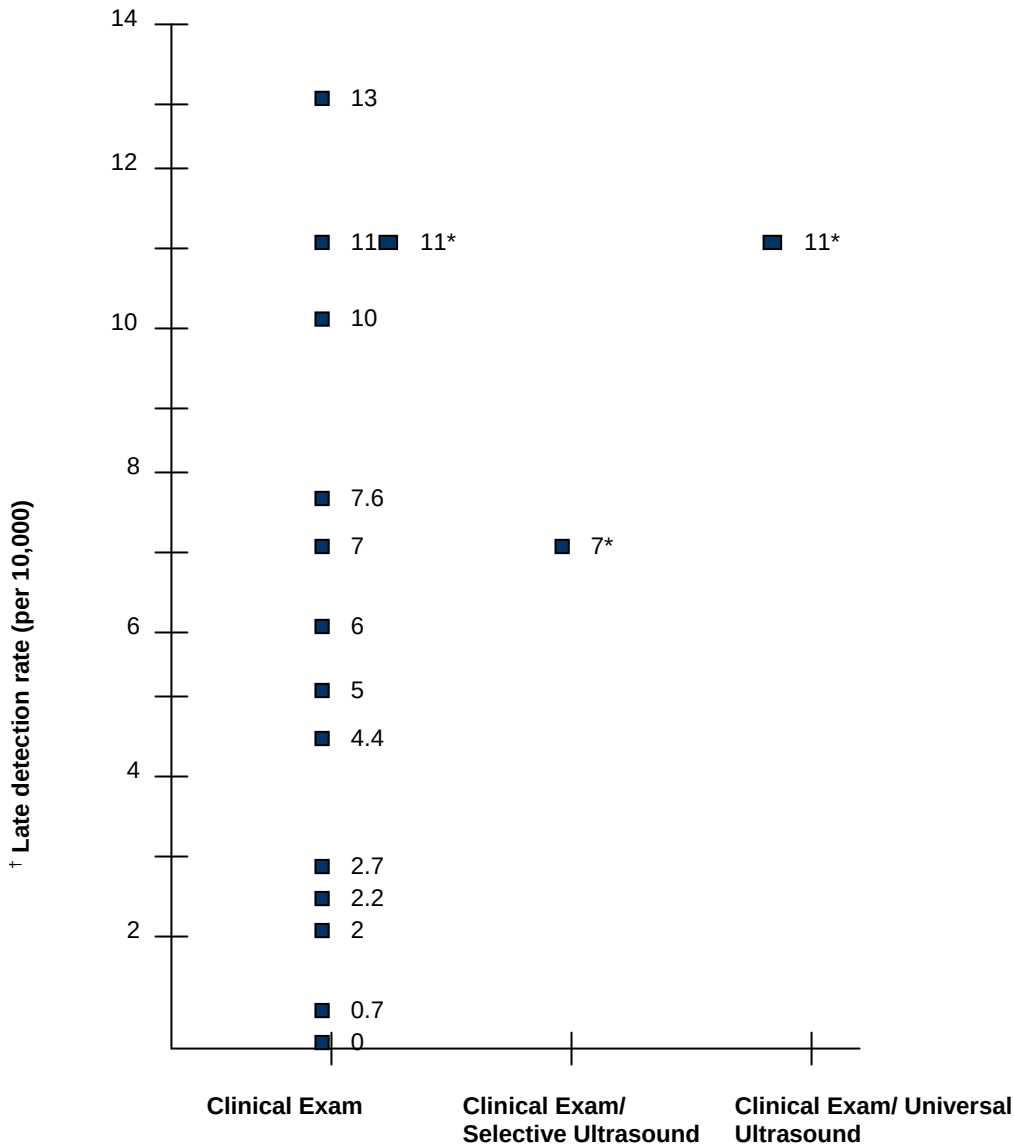
KQ 5: Does early diagnosis of DDH lead to early intervention, and does early intervention reduce the need for surgery or improve functional outcomes?

a) Is the likelihood of surgical intervention reduced in children diagnosed at an earlier age?

KQ 6: What are the adverse effects of early diagnosis and/or intervention?

KQ 7: What cost-effectiveness issues apply to screening for DDH?

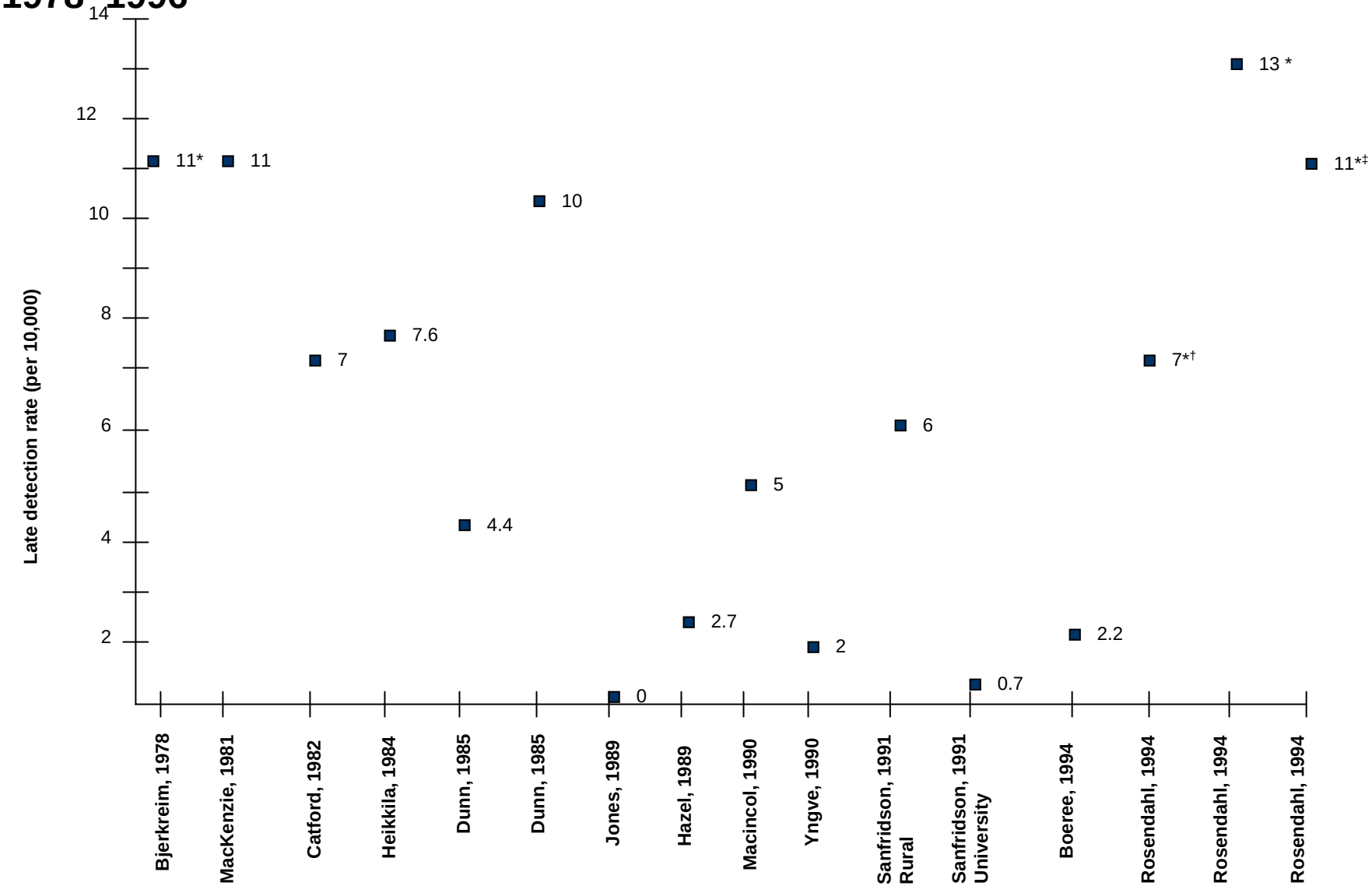
Figure 3. Variation in Late Detection Rate by Screening Method 1978–1996



*Subluxation and dislocation data only

†See Figure 4 for references

Figure 4. Variation in Late Detection Rate by Year of Study Publication 1978–1996



*Subluxation and dislocation only; †Clinical exam and selective ultrasound; ‡ Clinical exam and universal ultrasound

Figure 5. Range of Published Rates of Avascular Necrosis

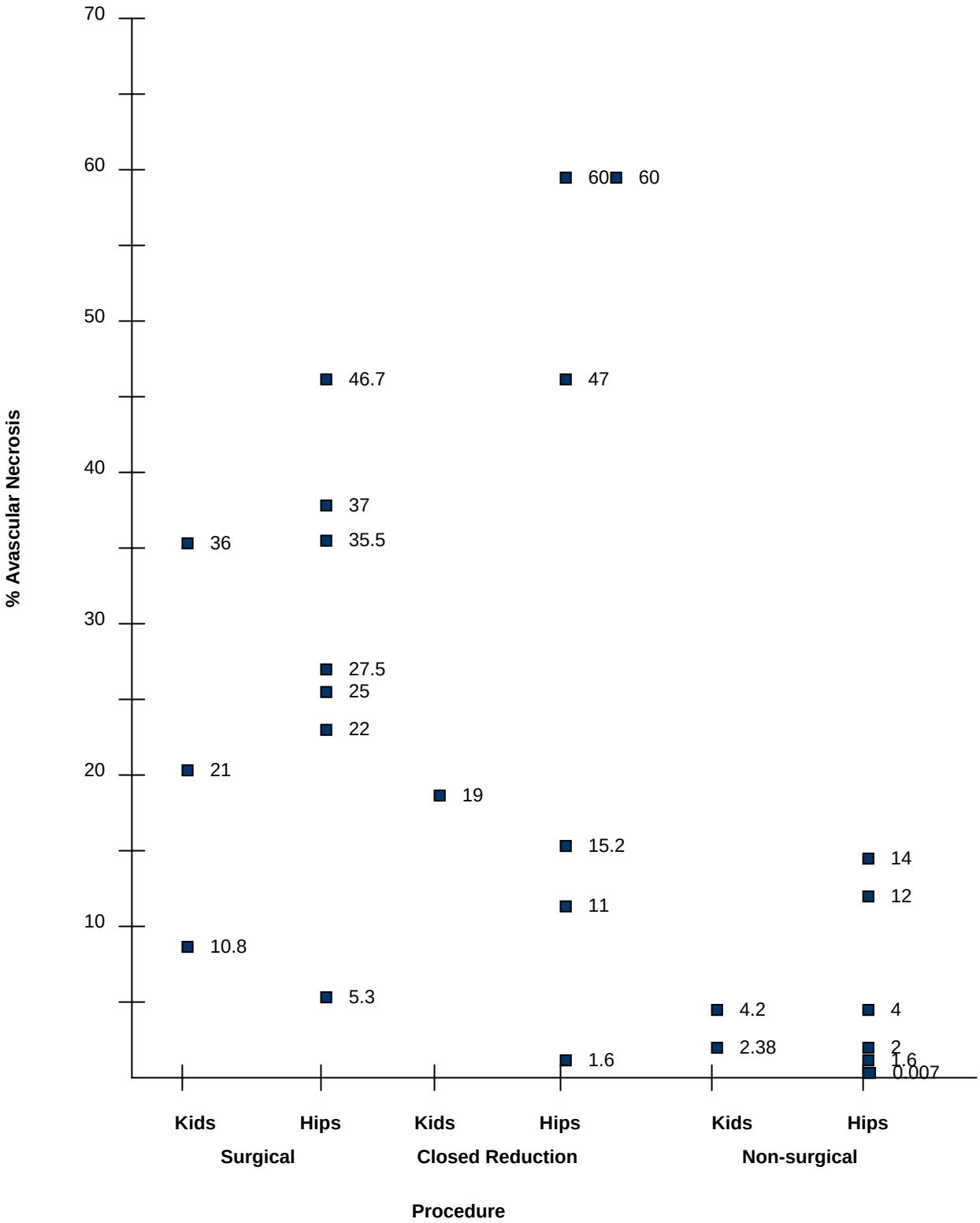


Figure 5. Avascular Necrosis Data Sources

Study, Year	%AVN	Condition	Type	Followup
Aksoy, 2002	15.2	CR	H	51 months
Broughham, 1990	47	CR	H	12 years (median)
Buchanan, 1981	36	S	K	8.3 years (mean)
Cashman, 2002	1.6	NS	H	6.5 years
Danielsson, 2000	5.3	S	H	10.9 years
Eidelman, 2003	0.007	NS	H	< 1 year
Grill, 1988	2.38	NS	K	4.46 years
Konigsberg, 2003	27.5	S	H	10.3 years
Kruczynski, 1996	14	NS	H	NR
Lennox, 1993	21	S	K	>1 year
Malvitz, 1994	60	CR	H	30 years
Pool, 1986	2	NS	H	>1 year
Pool, 1986	60	CR	H	>1 year
Powell, 1986	25	OR	H	51 months
Powell, 1986	22	OR	H	51 months
Powell, 1986	46.7	OR	H	51 months
Sosna, 1992	35.5	OR	H	11 years
Suzuki, 2000	4	NS	H	NR
Suzuki, 2000	12	NS	H	NR
Thomas, 1989	37	OR	H	9 years
Tumer, 1997	10.8	OR	K	8.1 years
Weiner, 1980	11	CR	H	NR
Yamada, 2003	1.6	CR	H	NR
Yoshitaka, 2001	19	CR	K	18 years
Yoshitaka, 2001	4.2	NS	K	18 years
AVN, avascular necrosis; CR, closed reduction; H, hips; K, kids;				
NS, non-surgical; OR, open reduction; S, surgical				

Table 1. Previous Recommendations by the American Academy of Pediatrics and the Canadian Task Force for DDH for Newborns and Infants*

Screening Interventions	Summary Findings	Level of Evidence**	Quality of Evidence	Recommendations
Serial clinical examination of the hips by a trained clinician during the periodic health examination in **all infants** (until walking independently)	Decreases the operative rate from 1–2 per 1000 infants to 0.2–0.7 per 1000	III	CTF: “Fair” AAP: "Good"	BOTH RECOMMEND CTF: (B) AAP: "Strong"
Ultrasound screening (static or dynamic method) in **all infants**	Increases splint rate without decreasing late surgery rate	II-1, III	Both: "Fair"	BOTH RECOMMEND AGAINST CTF: (D) AAP: "Strong"
Ultrasound screening in **high-risk infants**	Does not affect many infants and does not reduce the operative rate	II-1, III	CTF: “Fair” AAP: “Strong”	CTF RECOMMENDS AGAINST (D) AAP RECOMMENDS “Strong”
Routine radiographic screening at ages 3 to 5 months	Test is unreliable	III	CTF: “Fair”	CTF RECOMMENDS AGAINST (D)
Ultrasound or radiograph if positive clinical exam	Does not influence management		AAP: “Poor”	AAP RECOMMENDS AGAINST “Strong”
Positive screen: refer to orthopedist Equivocal screen: follow-up in 2 weeks Negative screen: follow-up in 2 months			AAP: "Good"	AAP RECOMMENDS "Strong"
Suspicious physical, not positive: refer to orthopedist or obtain ultrasound 3-4 weeks			None	AAP was "Mixed"
Treatment Interventions				
Triple diapering	Effectiveness unknown, may delay definitive therapy, though may aid in ensuring follow-up		AAP: “Poor”	AAP RECOMMENDS AGAINST “Strong”
Abduction therapy	Effectiveness unknown. Causes avascular necrosis of the hip (in 1%–4% of treated infants)	III		CTF DOES NOT RECOMMEND (C)
Observation before intervening	DDH resolves spontaneously in many cases	I	CTF: “Good”	CTF RECOMMENDS (A)

“I”, controlled trial with randomization; “II-1”, controlled trial without randomization; “III”, expert opinion; AAP – American Academy of Pediatrics; CTF – Canadian Task Force

*For more details on quality and recommendation coding, please see Patel (2001) & Lehmenn (2000).

Table 2. Risk Factors

Author, Year	N Overall	N with DDH	Risk Factor	Relative Risk	Patients with Risk Factor who Have DDH (%)	Number of DDH Positive Cases with Risk Factor (%)	Quality Rating
Andersson, 2001[33]	6,571	78 D or I*	Breech	3.72	D or I: 3.89%	D or I : 12.8%	Fair
		13 Treated		11.08	Treated: 1.56%	Treated: 30%	
Artz, 1975[34]	23408	312	Breech	6.35	6.64%	22.10%	Fair
			First born	1.31	1.71%	68%	
			Female	4.15	2.15%	79.50%	
Bache, 2002[35]	29,323	2340 92 treated	Breech	1.95, 4.14†	7.8%, 1.3%†	27%	Good
			Family history	3.4, 3.8†	13.4%, 1.2%†	7.60%	
			Female	1.7, 1.9†	6.6%, 0.59%†	91%	
			Breech female	2.8, 6.6†	11.0%, 2.0%†	14%	
			Family history and female	5.1, 3.7†	20.2%, 1.2%†	2.2%	
			Birth weight >4 kg	1.6, 1.8†	6.1%, .54%†	NR	
Boere-Boonekamp, 1998[37]	1,968	72	Breech	1.35	5.00%	4.2%	Fair
			Family history	2.59	9.6%	11.1%	
Boeree, 1994[36]	26,952	118	Breech	6.98	3.0%	10.2%	Fair
			Family history	24.9	10.7%	20.8%	
			Foot deformity	4.42	1.90%	2.5%	

Goss, 2002[38]	5,166	100	Breech	5.2	10.1%	24%	Fair
			Family history	NR	NR	25%	
			Female	3.3	6.4%	77%	
Holen, 1996[39]	408	25	Breech	5.55	6.1%	NR	Fair
Jones, 1989[40]	3,289	51	Breech	4.97	7.7%	11.8%	Fair
			Family history	10.8	16.7%	5.9%	
Miranda, 1988[41]	49,937	317	Breech	4.72	NR	17.4%	Fair
			First born	1.29	NR	53.0%	
			Female	1.67	NR	81.1%	
Sahin, 2004[42]	5,798	10	Breech	NA	<1% (1/111) overall	10% overall	Fair
			Family history	NA			
			Muscle/skeletal deformity	NA			
			Swaddling	NA			
Walter, 1992[43]	1,772	8	Breech	8.24 overall	4.12% overall	5% overall	Fair
			Family history				
			Postural abnormalties				
			Oligohydraminos				

D or I, dislocated or dislocatable; NR, not report
† = Ultrasound positive, treated

Table 3. Population-based Screening

Author, Year	Screening Technique	N	Information About Clinical Examiners	Reference Standard for DDH	When was Reference Standard Applied?	Clinical Exam Instability Rate/1,000 Children
Bache, 2002[35]	Clinical exam, universal ultrasound	29,323	Not specified	Need for treatment	Serial assessment over the first 6 weeks of life	NR
Bialik, 1998[14]	Clinical exam, universal ultrasound	4,321	Experienced neonatologist (number not specified)	Need for treatment	Serial assessment over the first 6 weeks of life	15.2
Giannakopoulou, 2002[16]	Ultrasound of clinical exam or risk factor +	6,140	2 experienced pediatricians	Ultrasound abnormality	15 days of age	17.9
Paton, 1999[44]	Ultrasound of clinical exam or risk factor +	20,452	Pediatrician (number not specified)	Dislocation on ultrasound	Within 2 weeks of birth for exam +; 8-9 weeks for risk factors +	14
Riboni, 2003[45]	Clinical exam, universal ultrasound	8,896	Neonatologists (number not specified)	Ultrasound abnormality	5 days of age	2.1 (only included frankly positive Barlow Ortolani tests)
Rosenberg, 1998[46]	Clinical exam, universal ultrasound	9,199	Experienced neonatologist (number not specified)	Unstable hips based upon clinical exam or ultrasound	During newborn hospitalization	14.5
Rosendahl, 1996[47]	Clinical exam, universal ultrasound	3,613	8 physicians with 2 or more years pediatrics experience	Dislocatable or dislocated on exam or "major" dyplasia on ultrasound	Within 24 hours of clinical exam	19.1

Table 3. Population-based Screening, Continued

Author, Year	Clinical Exam + Risk Factor Positive Rate/1,000 Children	% With DDH Identified Only by Exam	% Exam Positive Without DDH by Ultrasound	Initial Ultrasound Positive Rate/1,000 Children	% with DDH Identified Only By Ultrasound	Treatment Rate/1,000 Children	Late Diagnosis Rate/1,000 Children*
Bache, 2002[35]	NR	0%	18%	65.9 (hips) all abnormal ultrasound 39 for subluxable/dislocated	65%	3.1 (hips)	0
Bialik, 1998[14]	NR	0%	2%	55.3	52%	6.2 (hips)	NR
Giannakopoulou, 2002[16]	35.8	NA	41%	12.2	32%	10.6	NR
Paton, 1999[44]	54.1	NA	87%	1.8	31%	NR	0.4 dislocations
Riboni, 2003[45]	NR	NA	58%	28	56%	3.8	2.1 DDH 0.6 more severe than dysplasia
Rosenberg, 1998[46]	NR	5%	NA	68.2	50%	NR	NR
Rosendahl, 1996[47]	NR	11%	38%	29.6 static; 30.4 dynamic; 23.8 dislocatable/dislocated	28%	34	0.2 subluxation/ dislocation

NA, not applicable; NR, not reported
*Late presentation variably defined as after 1 month, 3 months, or 6 months in these studies.

Table 3. Population-based Screening, Continued

Author, Year	Rate/Timing of Spontaneous Resolution	Followup of Initially Negative Tests	Quality Rating	Comments
Bache, 2002[35]	96% of hips with ultrasound abnormalities at birth by 6 weeks	Unclear	Fair	Exam data only reported in relation to those ultimately requiring treatment
Bialik, 1998[14]	90.3% of hips with dysplasia or instability by 6 weeks	NR	Fair	
Giannakopoulou, 2002[16]	10/75 hips (10/10 with physiological dysplasia) within 4 weeks	NR	Fair	Risk factors not specified
Paton, 1999[44]	NR	Unclear	Fair	
Riboni, 2003[45]	206/215 with borderline dyplasia by 1 month	Ultrasound at three months (7,361/8,852): 19 additional cases identified and treated	Fair	3 patients with abnormal ultrasound did not have intervention despite recommendation: all healed normally Kappa for ultrasound and exam: 0.21
Rosenberg, 1998[46]	NR	NR	Fair	
Rosendahl, 1996[47]	13/16 with minor dysplasia by 1-2 months	Unclear		Kappa for hip stability ultrasound and exam: 0.6

Table 4. Randomized Controlled Trials of Screening

Author, Year	Risk Factors	Screening Approach	Adequate Randomization/ Allocation Concealment	Screening Tests	Screening Clearly Defined	# Analyzed/ # Available	Screened Positive Per Thousand
Elbourne, 2002[48] Group 1	NR	Group 1: ultrasound	Centralized randomization service; participant characteristics listed but statistical comparison not reported	Ultrasound (dynamic)	Yes	258/314	385 (38.5%)
Elbourne, 2002[48] Group 2		Group 2: clinical assessment only		Clinical exam (otherwise not specified)		276/315	492 (49.2%)
Holen, 2002[49] Group 1	Gender, family history, breech, hip instability on examination, doubtful clinical findings on Barlow test, foot deformities	Group 1: clinical exam, ultrasound	Random numbers table; no allocation concealment; participant characteristics listed but not statistical comparison not reported	Ultrasound (dynamic), Barlow and Ortolani tests	Yes	7,489/7,840	9.6

Holen, 2002[49] Group 2		Group 2: clinical exam, selective ultrasound		Barlow and Ortolani tests, and dynamic ultrasound if risk factor +		NR/7689	8.6
Rosendahl, 1994[21] Group 1	Family history, breech presentation, hip instability	Group 1: general ultrasound	Patients assigned based upon hospital unit to which they were admitted; some risk factors not equally distributed	Clinical exam and ultrasound (static and dynamic)	Yes	NR/3613	Barlow/Ortolani +: 16Any risk factor +: 126Requiring treatment: 34
Rosendahl, 1994[21] Group 2		Group 2: selective ultrasound		Clinical exam with ultrasound (static and dynamic) for high risk		NR/4388	Barlow/Ortolani +: 15 Any risk factor +: 118 Requiring treatment: 20
Rosendahl, 1994[21] Group 3		Group 3: clinical assessment only		Clinical exam (Barlow/Ortolani tests)		NR/3924	Barlow/Ortolani +: 18 Any risk factor +: NR Requiring treatment: 18

AVN, avascular necrosis; NR, not reported; NA, not available

Table 4. Randomized Controlled Trials of Screening, Continued

Author, Year	Lost Cases	Handling of Drop-outs/ Cross-overs	Followup Duration and Adequacy	Required Further Treatment	Adverse Effects	Late Detected Hip Dysplasia	Quality Rating
Elbourne, 2002[48] Group 1	56 (incomplete data)	Drop-outs excluded from analysis; cross-overs analyzed in an intention-to-treat fashion	Until 2 years of age maximum marginally adequate follow-up to ensure identification; inadequate for AVN identification & functional outcomes; followup standardized for first 8 weeks, then care per "usual pattern" at each of 33 sites	17	9 AVN	NR 4% or 23 cases had late treatment, and "both groups had much the same proportions with late treatment"	Fair
Elbourne, 2002[48] Group 2	39 (incomplete data)			20	7 AVN	NR (see above)	
Holen, 2002[49] Group 1	351 (NICU due to low birth weight)	Drop-outs excluded from analysis; no data provided on cross-overs	6-11 years (mean: 8.5 years); duration sufficient to identify late cases, may not be adequate to identify AVN, nor to assess functional outcomes; followup protocol not described in detail	2	0	1 late case: .13 per 1000 (8 years)	Fair

Holen, 2002[49] Group 2	None reported			1	1 AVN	5 late cases: .65 per 1,000 (8 years); 1 closed reduction, 2 abduction splint, all female	
Rosendahl, 1994[21] Group 1	34 missing exam data, 40 missing risk factor data	NR	27 months minimum (mean: 42 months); probably adequate duration to identify late cases, but not to identify all with AVN nor functional outcomes; followup protocol not described in detail	1 abduction splintuntreated but req. followup due to screen +: 130/1,000	NR	5 late cases (4 dysplasia, 1 subluxation),Rate: 1.4/1,000	Fair
Rosendahl, 1994[21] Group 2	36 missing exam data, 44 missing risk factor data			2 abduction splint, 1 surgery untreated but req. followup due to screen +: 18/1000	NR	9 late cases (6 dysplasia, 2 subluxation, 1 dislocation), Rate: 2.1/1,000	
Rosendahl, 1994[21] Group 3	None reported			3 abduction splint, 2 surgery untreated but req. followup: NA	NR	10 late cases (5 dysplasia, 3 subluxation, 2 dislocation), rate: 2.6/1,000	

Table 5. Randomized Controlled Trial Training Approaches

Study, Year, Setting	Clinical Exam Only				Selective Ultrasound				Universal Ultrasound			
Study, Year Setting	**N**	**Treatment Rate***	**Late Diagnosis Rate***	**Subsequent Treatment Required*†**	**N**	**Treatment Rate***	**Late Diagnosis Rate***	**Subsequent Treatment Required*†**	**N**	**Treatment Rate***	**Late Diagnosis Rate***	**Subsequent Treatment Required*†**
Elbourne, 2002[48] Infants <43 days of age, with hip instability on initial exam, referred to 33 clinics, UK 1994-98	315	492	NR	63					314	385	NR	54
Holen, 2002[49] Newborns 1-3 days old in hospital, Norway 1988-92					7,689	8.6	0.65	0.13	7,640	9.6	0.13	0.26
Rosendahl, 1994[21] Newborns in maternity hospital, Norway 1988-90	3,924	18	2.6	1.2	4388	20	2.1	0.7	3,618	34	1.4	0.3

NR, not reported
*Number per 1,000 children
†Additional treatment required, indicating primary treatment unsuccessful

Table 6. Interventions

Author, Year	Years Inclusive	Diagnostic Standard	Surgical or Nonsurgical Invervention	Previous Intervention	Patient Age in Months at Intervention: Average (Range)	Original N
Aksoy, 2002[92]	NR	Any ultrasound abnormality (including mild dysplasia, subluxation, dislocation)	Surgery (traction ranging 5-45 days followed by closed reduction +/- adductor tenotomy), followed by immobilization for 45-190 days	NR	6 (2-13)	129 patients with 200 treated hips
Cashman, 2002[93]	1988 - 1997	Dislocation or displacement on ultrasound	Nonsurgical (Pavlik harness)	None	NR	332 patients with 546 treated hips
Danielsson, 2000[94]	1977-1991	Radiographically definite dislocation or subluxation, clinically dislocated or dislocatable	Surgery (adductor tenotomy followed by 3-4 wks traction, then closed reduction), followed by hip spica for 6 months, then Pavlik harness continuously for 3 months, then night only for 3 months	NR	10 (2-64)	71
Eidelman, 2003[95]	1992 - 2001	Radiographic and/or clinical instability persisting 2 or more weeks	Nonsurgical (Pavlik's method)	None	8.5 weeks (5-12)	75 patients with 127 treated hips
Konigsberg, 2003[96]	1981-1997	NR	Surgery (open reduction), followed by hip spica cast (duration range: 6 to 25 wks), followed by abduction bracing for "varying periods"	Pavlik in 24 pts, closed reduction in 5, preop traction in 27	7.7 (2.4-18.9)	32 patients with 40 treated hips

Author, Year	Duration of Followup in Years: Average (Range)	Loss to Followup	Rate of Success (Reduction of Hip, Not Requiring Further Intervention)	Rate of Avascular Necrosis	Functional Outcomes Assessed
Aksoy, 2002[92]	4.3 (1.3-20)	39/200 hips	76% good results radiographically; 82% good results based upon functional/clinical findings Subsequent surgeries NR	25/164 (15.2%) among treated hips with "sufficient radiographs"	Categorized by amount of pain, stability of hip, range of motion, and Trendelenberg's sign: 82% of patients had good results, 18% poor results. not standardized
Cashman, 2002[93]	6.5 (2.1-11.8)	37/332 (11.1%)	18 hips (16 patients) required subsequent surgery	3/316 patients	NR
Danielsson, 2000[94]	10.9 (5.5-17.5)	None	54/71 (76%) did not require open surgery	4/75 treated hips (5.3%)	2 patients with pain, both later avoided some activities. not standardized
Eidelman, 2003[95]	1 (range not given)	NR	3/75 patients required closed reduction and spica casting	1/127 treated hips	NR
Konigsberg, 2003[96]	10.3 (2.5-18.6)	None (see comments) 14/32 patients had gait analyzed; 20/32 patients had assessment for pain	23/32 did not require further surgery	11/40 treated hips (27.5%)	14 patients old enough for gait analysis with range of motion and strength testing: 12/14 with less than 5% limitation in dynamic range of motion; 20 patients received a standardized screen for pain: 8/20 positive, leading to activity limitation in 2

Author, Year	Outcome Assessment Blinded	Age Effects	Relative Quality Rating	Comments
Aksoy, 2002[92]	No	101/106 hips treated at <6 mos of age had "satisfactory outcome" radiographically and functionally, versus 75/83 hips in children >6 mos of age	Fair-Poor	Retrospective case series, not clearly a consecutive series of patients; number of surgeons NR; rationale for variation in type/duration of therapy NR; identified a strong association between AVN and worse functional outcomes
Cashman, 2002[93]	No	None reported/obtainable	Fair-Poor	Prospective case series; patients presenting at > 90 days of age or with previous treatment were excluded; cases managed by one surgeon; therapeutic approach not detailed
Danielsson, 2000[94]	No	AVN in 1/36 pts < 6 mos; 3/17 in pts 6 - 11 mos; 0/13 pts 12-22 mos. Subsequent surgery required in 1/37 hips < 6mos; 3/17 hips 6-11 mos; 7/13 hips 12-22 mos	Fair	Prospective consecutive observational cohort; single surgeon involved in all but 1 case; standard therapy applied in all cases; authors note limited followup and express belief that "further clinical and radiographic deterioration can be expected in the long run"
Eidelman, 2003[95]	No	Age when treatment started had no effect on duration of treatment; those requiring surgical followup were in youngest age group (3 weeks)	Poor	Retrospective case series; excludes approximately 30% of patients who were screened positive but treated elsewhere; number of treating physicians NR; limited follow-up period
Konigsberg, 2003[96]	No	AVN less common in children < 6 months (1/20 hips vs 10/20 hips >6 months of age); mean age without AVN 6 months, mean age with AVN 11 months; those reporting pain had mean age of 14 years vs. 11 years with no pain	Fair	Retrospective case series, limited to patients with >2 yrs followup; number of surgeons NR; rationale for variation in type/duration of therapy NR; authors note that many patients have not reached skeletal maturity, so more may require surgery; association between age and AVN, age and pain not controlled for any confounders

Table 6. Interventions, Continued

Author, Year	Years Inclusive	Diagnostic Standard	Surgical or Nonsurgical Invervention	Previous Intervention	Patient Age in Months at Intervention: Average (Range)	Original N
Gregersen, 1969[97]	1958-1966	Ortolani positive or unstable on clinical exam	Nonsurgical (plaster casting in abduction) for 13 weeks, followed by abduction therapy in patients with persistent subluxation	None	6 days (no range given); all before 19 days	59 patients with 81 treated hips
Malvitz, 1994[98]	1940-1969	Ortolani positive	Surgery (closed reduction) followed by spica cast for 3 months, followed by full-time abduction brace and subsequent night-only brace for "several years"	NR	21 (1-96)	119 patients with 152 treated hips
O'Hara, 1988[99]	5 years (not specified)	Radiographic dislocation or instability on arthrogram	Surgery (open reduction) followed by immobility in spica cast for 3 months, then cast modified to allow movement for another 1.5-3 months	9 with abduction splinting, 4 with double diapering	6.5 (3-15)	40 patients with 40 treated hips
Sosna, 1992[100]	1970-1985	NR	Surgery (open reduction), no casting, followed by biomechanical device duration not specified	NR	9.3 (5-23)	70 patients with 78 treated hips
Tegnander, 2001[101]	1988 - 1990	Clinical instability and ultrasound abnormality in first week of life	Nonsurgical (Frejka pillow) duration 4 mos	None	Less than 7 days	108 patients with 144 treated hips

Author, Year	Duration of Followup in Years: Average (Range)	Loss to Followup	Rate of Success (Reduction of Hip, Not Requiring Further Intervention)	Rate of Avascular Necrosis	Functional Outcomes Assessed
Gregersen, 1969[97]	4.7 (1.6-8.7)	18/59 patients	40/41 patients (97.5%) good functional results (1 patient with re-dislocation)	NR	"No patient had any complaints"40/41 patients had normal gait39/41 had normal mobilityNot standardized
Malvitz, 1994[98]	30 (15-53)	None (see comments)	106/119 patients required no surgery	60% of treated hips had evidence of AVN	Used standardized scales for hip pain and functional use (gait and activities): 112/149 (75%) of hips had excellent outcome (94% of hips in patients <30 yrs of age, 57% of hips in patients >30 years of age)
O'Hara, 1988[99]	4.6 (2-7)	None	No patients required subsequent surgery	NR	Average age at walking: 14.1 months 6 patients with pain (5 mild without limited activity) 6 patients with minor limp not standardized
Sosna, 1992[100]	11 (5-18)	14/70 patients	47/62 (76%) hips had satisfactory clinical & radiographic results at final follow-up 53 patients had further surgery (reasons not discussed)	22/62 treated hips (35.5%)	Clinically satisfactory results in at least 76%: no pain, <50% restriction of range of motion, <1cm leg length reduction, and negative Trendelenberg's sign
Tegnander, 2001[101]	4.3 (1.5-6)	34/108 (21%)	3 patients (denominator NR) required further nonsurgical splinting/ casting	1 patient (denominator NR)	Thirteen patients noted to have intoeing gait (only two complained of this) not standardized

Author, Year	Outcome Assessment Blinded	Age Effects	Relative Quality Rating	Comments
Gregersen, 1969[97]	Yes, for radiographic findings only	The 32 hips <12 months of age at intervention had excellent functional results 94% of the time and fair/poor results 3%, versus the 120 hips >12 months with excellent 70% and fair/poor 19%; functional and radiographic results similar for patients <6 months and those 6-12 months	Fair	Retrospective case series; study excluded patients treated successfully with abduction splint, managed initially at another hospital, had open reduction as primary therapy, or if pre- or post-reduction radiographs were missing; of 154 remaining eligibles, 119 included in study; all management by one of 2 surgeons; study did not control for possible confounding of association of age at intervention with outcomes by age at final followup
Malvitz, 1994[98]	No	None reported/obtainable	Fair-Poor	Retrospective consecutive case series; excluded bilateral dislocations; number of surgeons NR
O'Hara, 1988[99]	No	None reported/obtainable	Fair-Poor	Retrospective case series, number of surgeons NR; previous nonsurgical therapy NR
Sosna, 1992[100]	No	None reported/obtainable	Fair-Poor	Retrospective case series, number of surgeons NR; previous nonsurgical therapy NR
Tegnander, 2001[101]	No	None reported/obtainable	Fair-Poor	Retrospective case series; inception cohort unclear; number of treating physicians NR; moderate loss to followup

Table 6. Interventions, Continued

Author, Year	Years Inclusive	Diagnostic Standard	Surgical or Nonsurgical Invervention	Previous Intervention	Patient Age in Months at Intervention: Average (Range)	Original N
Tumer, 1997[102]	NR	NR ("developmentally dislocated")	Surgery (open reduction with no preliminary traction), followed by hip spica for 3 months, then a night-only abduction splint for 3-4 months	Pavlik in 7 pts	11.2 (2-25)	37 patients with 56 treated hips
Yamada, 2003[103]	NR	NR ("dislocated")	Surgery (closed reduction) preceded by traction, followed by spica cast for 4 weeks, then abduction device for at least 6 months	"Most" treated previously with Pavlik harness	11.5 (6-23)	55 patients with 62 treated hips
Yoshitaka, 2001[104]	1963 - 1980	X-ray diagnosis of subluxation (no physical exam information provided)	Pavlik harness (213 patients): for 3 mos as of 1970 -- no information about duration from 1963 - 1969, nor # of patients in each timeframe Closed reduction (16 patients) followed by cast immobilization, duration unspecified	NR	4 months (1-24)	229 patients with 262 treated hips

Author, Year	Duration of Followup in Years: Average (Range)	Loss to Followup	Rate of Success (Reduction of Hip, Not Requiring Further Intervention)	Rate of Avascular Necrosis	Functional Outcomes Assessment
Tumer, 1997[102]	8.1 (3-17)	None	46/56 did not require further surgery (81%)	5/56 treated hips (9%)	1 patient with abnormal gait. not standardized
Yamada, 2003[103]	over 6 years	28/55 (52%)	11 of 31 hips required subsequent surgery	1/62 treated hips	NR
Yoshitaka, 2001[104]	19.1 (14-34)	None (see comments)	258/262 hips required no further intervention	12/262 treated hips (4.6%)	NR

Author, Year	Outcome Assessment Blinded	Age Effects	Relative Quality Rating	Comments
Tumer, 1997102	No	None reported/obtainable	Poor	Retrospective case series, not clearly a consecutive series of patients; single surgeon with uncertain number of surgical residents
Yamada, 2003103	No	Age at reduction no different among those with and without residual subluxation	Poor	Retrospective case series; inception cohort unclear; large loss to followup; number of surgeons NR
Yoshitaka, 2001104	No	NR	Fair-Poor	Retrospective case series; excluded from analysis 202/431 patients with subluxation due to loss to follow-up; no comparison of demographics, etc. of group excluded to those included; conducted non-blinded radiographic outcomes assessment only; number of surgeons NR

Table 7. Summary of Evidence

Arrow	Key Question	Level and Type of Evidence
1	Does screening for DDH lead to reduced need for surgery or improved functional outcomes?	Poor: no controlled studies have compared screening with no screening to determine whether there is an impact on functional outcomes; there is conflicting evidence from ecologic studies that screening reduces rates of surgery
2	Can infants at high risk for DDH be identified, and does this group warrant a different approach to screening than children at average risk?	Fair: in case-control and cohort studies, family history, breech presentation, and clinical instability are consistently associated with a high risk of DDH, but most infants with DDH do not have risk factors poor: practice-based, prospective studies on on the performance of risk assessment instruments are lacking
3	Does screening for DDH lead to early identification of children with DDH?	See 3a, 3b, 3c below
3a	What is the sensitivity, specificity, and predictive value of screening exams? (e.g., Barlow/Ortolani, other exam findings, ultrasonography, and radiographs)	Poor: ascertainment of test characteristics is unreliable, because definitions of a positive test vary, and most studies did not use an independent standard to determine disease status; low risk/ screen negative patients rare followed with intensity of high risk/screen positive patients; high rates of spontaneous resolution have been reported fair: most hip dysplasia identified by early ultrasound will resolve spontaneously in first weeks of life
3b	How does the age of the child affect screening parameters?	Fair: limited hip abduction becomes a more sensitive sign of DDH over the first several months of life

3c	How does the educational level and training of the screener impact screening?	Fair: experience with the clinical examination of the hip in infants predicts screen positive rates and accuracy of exam, but few head-to-head comparisons without biases have been conducted; consistent but limited amount of evidence that well-trained non-physicians can interpret clinical examination findings as well as pediatricians and better than physicians-in-training
4	What are the adverse effects of screening?	Poor: in theory, forceful exam of already-lax newborn hips might cause injury or dislocation, but there is limited and conflicting evidence regarding this hypothesis
5	Does early diagnosis of DDH lead to early intervention, and does early intervention lead to improved functional outcomes? Is the likelihood of surgical intervention reduced in children diagnosed at an earlier age?	Fair: early diagnosis leads to early intervention; evidence of the effectiveness of intervention is inconclusive, due to 1) high rate of spontaneous resolution, 2) absence of comparative studies of intervention vs. no intervention, 3) variation in surgical indications and protocols; few studies examine functional outcomes in a valid and reliable fashion fair-poor: evidence is limited and mixed on the effect of earlier diagnosis on likelihood of surgery
6	What are the adverse effects of early diagnosis and/or surgical and non-surgical interventions?	Fair: all nonsurgical and surgical interventions are associated with a risk of avascular necrosis; many nonsurgical interventions are in use, but data are insufficient to determine whether there are differences among them; this is also true of surgical interventions

Appendix 1. Devices and Procedures

Nonsurgical Abduction Devices

Abduction brace
Becker device
Craig splint
Divaricator splint
Frejka pillow
Immobilization in hip spica
Knee splint harness
Pavlik harness
Traction
von Rosen

Surgical Procedures

Anterolateral open reduction
Closed reduction
Dega osteotomy
Derotation osteotomy
Femoral osteotomy
Fergusen medial approach
Harris hip-rating system
Kalamchi modification of salter osteotomy
Ludloff's medial approach
Medial adductor open reduction
Medial open reduction
Open reduction
Osteotomy of the pelvis
Pemberton osteotomy
Pemberton acetabuloplasty
Salter Innominate osteotomy
Varus derotational osteotomy

Salvage Procedures

Chairi pelvic osteotomy
Colonna's operation
Rotational acetabular osteotomy
Shelf operation
Total hip arthroplasty
Triple innominate osteotomy

Appendix 2. Search Strategies

Database: MEDLINE (1966 to January 2005)

Screening and Adverse Effects of Screening

Screening

1 exp Hip Dislocation, Congenital/pa, di, ra, ri, us
2 exp "Diagnostic Techniques and Procedures"/
3 exp Hip Dislocation, Congenital/
4 2 and 3
5 1 or 4
6 exp "Sensitivity and Specificity"/
7 5 and 6
8 limit 7 to (English language and all child <0 to 18 years>)

Cost analysis

9 exp "Costs and Cost Analysis"/
10 5 and 9
11 exp Hip Dislocation, Congenital/ec
12 10 or 11
13 limit 12 to (English language and all child <0 to 18 years>)

Age of child and frequency of screening

14 exp Age Factors/
15 5 and 14
16 exp time factors/
17 5 and 16
18 15 or 17
19 limit 18 to (English language and all child <0 to 18 years>)

Risk of DDH

20 exp RISK/
21 5 and 20
22 limit 21 to (English language and all child <0 to 18 years>)

Diagnostic errors of screening

23 exp Diagnostic Errors/
24 5 and 23
25 limit 24 to (English language and all child <0 to 18

Educational level and training of screener

26 family physician$.mp. or exp Physicians, Family/
27 primary care.mp. or exp Primary Health Care/exp NURSES/ or exp NURSE'S ROLE/ or exp
28 NURSES' AIDES/ or exp nursing care/

29 exp Pediatrics/ or pediatrician$.mp.
30 exp physician assistants/ or physican assistant$.mp.
31 26 or 27 or 28 or 29 or 30
32 5 and 31
33 limit 32 to (English language and all child <0 to 18 years>)

Adverse effects of screening
34 (adverse$ adj5 effect$).mp.
35 5 and 34
36 exp "Wounds and Injuries"/et
37 5 and 36
38 35 or 37
39 limit 38 to (English language and all child <0 to 18 years>)

Types of studies
40 Comparative Study/
41 exp Evaluation Studies/
42 exp Epidemiologic Studies/
43 40 or 41 or 42
44 5 and 43
45 limit 44 to (English language and all child <0 to 18

Age limit
46 limit 44 to (English language and all infant <birth to 23 months>)

Barlow
1 barlow.tw.
2 ortolani.mp.
3 1 or 2
4 (hip or dysplas$ or cdh or congenit$).mp.
5 3 and 4
6 limit 5 to human
7 limit 6 to english language
8 limit 6 to abstracts
9 7 or 8
10 from 9 keep 1-64

Outcomes, Interventions, and Adverse Effects of Interventions

Outcomes
1 exp Hip Dislocation, Congenital/nu, pc, dt, rh, su, th
2 exp "Outcome and Process Assessment (Health Care)"/
3 1 and 2
4 limit 3 to (English language and all child <0 to 18 years>)

Age of child

5 exp Age Factors/
6 1 and 5
7 limit 6 to (English language and all child <0 to 18 years>)

Types of studies

8 Comparative Study/
9 exp Evaluation Studies/
10 exp Epidemiologic Studies/
11 8 or 9 or 10
12 1 and 11
13 limit 12 to (English language and all child <0 to 18 years>)

Surgical vs. non-surgical

14 exp Hip Dislocation, Congenital/nu, pc, dt, rh, th
15 2 and 14
16 limit 15 to (English language and all child <0 to 18 years>)
17 exp Hip Dislocation, Congenital/su
18 2 and 17
19 limit 17 to (English language and all child <0 to 18 years>)

Time factors, diagnoses, and outcomes

20 exp Time Factors/
21 exp Hip Dislocation, Congenital/pa, di, ra, ri, us
22 exp Hip Dislocation, Congenital/
23 exp diagnosis/
24 22 and 23
25 21 or 24
26 17 and (5 or 20) and 25
27 limit 26 to (English language and all child <0 to 18 years>)

Adverse effects of interventions

28 (adverse$ adj5 effect$).mp.
29 1 and 28
30 exp "Wounds and Injuries"/et
31 1 and 30
32 29 or 31
33 limit 32 to (English language and all child <0 to 18 years>)

Pavlik

1 pavlik.tw.
2 (hip or dysplas$ or cdh or congenit$).mp.
3 1 and 2
4 limit 3 to human
5 limit 4 to english language

6 limit 4 to abstracts
7 5 or 6
8 from 7 keep 1-129

Classification and Differential Diagnosis of DDH

1 exp Hip Dislocation, Congenital/cl [Classification]
2 exp Hip Dislocation, Congenital/
3 diagnosis, differential/
4 2 and 3
5 1 or 4
6 limit 5 to (English language and all child <0 to 18 years>)

Cochrane Databases

Searched the keyword “hip dysplasia” through June 2004 in the following databases:

- Cochrane Central Register of Controlled Trials
- Cochrane Database of Systematic Reviews
- Cochrane DSR, ACP Journal Club, and DARE

Automatic Updates

Weekly updates for key question’s 2,3,4 combined
Weekly updates for key question’s 2,3,4,5,6 combined

Appendix 3. Abstract Inclusion/Exclusion Criteria

Developmental Dysplasia of the Hip

Include	Randomized control trial Comparative study (cohort, case-control, or observational) Systematic reviews, standards, or guidelines Relevant for risk Relevant for screening Relevant for harms Relevant for intervention Relevant for harms of intervention Relevant for cost Relevant for background or epidemiology Relevant for methodology US or non-US (if applicable) population
Exclude	Letters, editorials, nonsystematic reviews, commentaries Case reports Less than 5 cases Screened age beyond six months Intervention age (mean) beyond one year Children with teratological DDH (neuromuscular disorders, multiple congenital anomalies) Salvage procedures Information not relevant or outside scope Poor quality study Non-English language Non-human

Appendix 4. Quality Rating Criteria

Diagnostic Accuracy Studies

Criteria

- Screening test relevant, available for primary care, adequately described
- Study uses a credible reference standard, performed regardless of test results
- Reference standard interpreted independently of screening test
- Handles indeterminate results in a reasonable manner
- Spectrum of patients included in study
- Sample size
- Administration of reliable screening test

Definition of ratings based on above criteria

Good: Evaluates relevant available screening test; uses a credible reference standard; interprets reference standard independently of screening test; reliability of test assessed; has few or handles indeterminate results in a reasonable manner; includes large number (more than 100) broad-spectrum patients with and without disease.

Fair: Evaluates relevant available screening test; uses reasonable although not best standard; interprets reference standard independent of screening test; moderate sample size (50 to 100 subjects) and a "medium" spectrum of patients.

Poor: Has important limitations such as: uses inappropriate reference standard; screening test improperly administered; biased ascertainment of reference standard; very small sample size of very narrow selected spectrum of patients.

Randomized Controlled Trials (RCTs) and Cohort Studies

Criteria

- Initial assembly of comparable groups**:** RCTs—adequate randomization, including concealment and whether potential confounders were distributed equally among groups; cohort studies—consideration of potential confounders with either restriction or measurement for adjustment in the analysis; consideration of inception cohorts
- Maintenance of comparable groups (includes attrition, cross-overs, adherence, contamination)
- Important differential loss to follow-up or overall high loss to follow-up
- Measurements: equal, reliable, and valid (includes masking of outcome assessment)
- Clear definition of interventions
- Important outcomes considered
- Analysis: adjustment for potential confounders for cohort studies, or intention-to-treat analysis for RCTs

Definition of ratings based on above criteria

Good: Meets all criteria: Comparable groups are assembled initially and maintained throughout the study (follow-up at least 80 percent); reliable and valid measurement instruments are used and applied equally to the groups; interventions are spelled out clearly; important outcomes are considered; and appropriate attention to confounders in analysis.

Fair: Studies will be graded "fair" if any or all of the following problems occur, without the important limitations noted in the "poor" category below: Generally comparable groups are assembled initially but some question remains whether some (although not major) differences occurred in follow-up; measurement instruments are acceptable (although not the best) and generally applied equally; some but not all important outcomes are considered; and some but not all potential confounders are accounted for.

Poor: Studies will be graded "poor" if any of the following major limitations exists: Groups assembled initially are not close to being comparable or maintained throughout the study; unreliable or invalid measurement instruments are used or not applied at all equally among groups (including not masking outcome assessment); and key confounders are given little or no attention.

Case Control Studies

Criteria

- Accurate ascertainment of cases
- Nonbiased selection of cases/controls with exclusion criteria applied equally to both
- Response rate
- Diagnostic testing procedures applied equally to each group
- Measurement of exposure accurate and applied equally to each group
- Appropriate attention to potential confounding variable

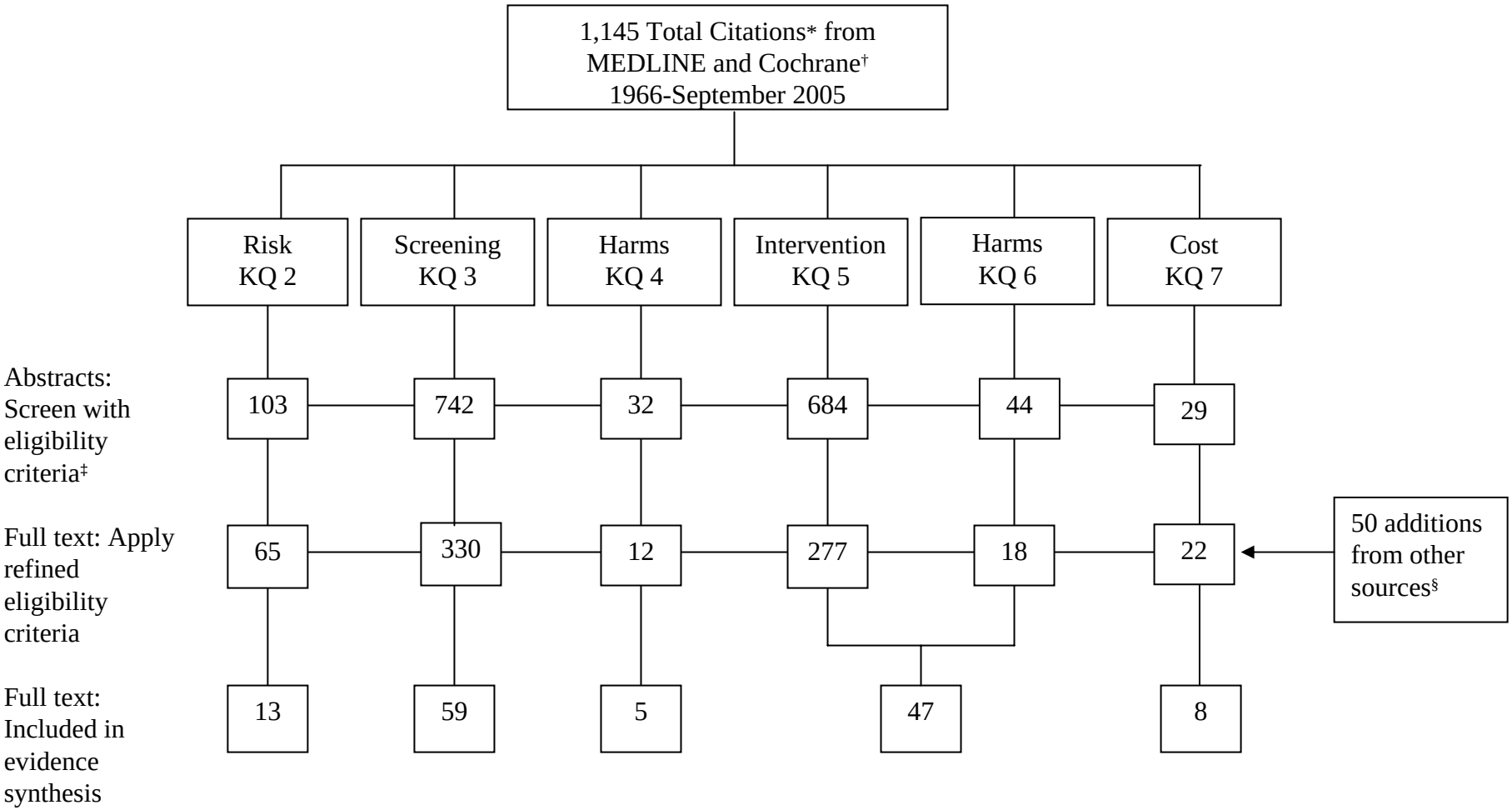
Definition of ratings based on above criteria

Good: Appropriate ascertainment of cases and nonbiased selection of case and control participants; exclusion criteria applied equally to cases and controls; response rate equal to or greater than 80 percent; diagnostic procedures and measurements accurate and applied equally to cases and controls; and appropriate attention to confounding variables.

Fair: Recent, relevant, without major apparent selection or diagnostic work-up bias but with response rate less than 80 percent or attention to some but not all important confounding variables.

Poor: Major selection or diagnostic work-up biases, response rates less than 50 percent, or inattention to confounding variables.

Appendix 5. Search and Selection of the Literature



*English language only

†Cochrane Databases include the Cochrane Central Register of Controlled Trials, Cochrane Database of Systematic Reviews, Cochrane DSR, ACP Journal Club, and DARE

‡ Duplicates may exist between key questions at all three search and selection levels

§ Other sources include reference lists, expert referrals, etc.

Appendix 6. Expert Reviewers

Content Experts
Viktor Bialik, MD, PhD
Director, Pediatric Orthopedics Unit
Rambam Medical Center/ Technion-Israel Institute of Haifa, Israel

Michael Goldberg, MD
Chairman, Department of Orthopedics
Tufts University School of Medicine

Harold Lehmann, MD, PhD
Associate Professor, Department of Pediatrics and Health Sciences Informatics
Johns Hopkins Medical Institutions

Hema Patel, MD
Montreal Children's Hospital /McGill University Health Centre
Montreal, Quebec

Ronald Turker, MD
Head, Pediatric Orthopaedics
Oregon Health and Science University

Professional Societies or Organizations
Jon R Davids, MD
Assistant Chief, Shriners Hospitals for Children and Director, Motion Analysis Lab
Orthopaedic Surgery Faculty
On behalf of the Pediatric Orthopaedic Society of North America (POSNA)

Michael Vitale, MD, MPH
Assistant Professor, Division of Pediatric Orthopaedic Surgery and Assistant Professor of Orthopaedic Surgery / Columbia University
On behalf of the American Academy of Orthopaedic Surgeons/American Association of Orthopaedic Surgeons (AAOS), the Pediatric Orthopaedic Society of North America (POSNA), and the American Academy of Pediatrics (AAP)

US Preventive Service Task Force Leads
Carol Loveland-Cherry, PhD, RN, FAAN

Virginia Moyer, MD, MPH

Barbara Yawn, MD, MSc, FAAFP

Keith R Gabriel, MD

Federal Agencies
Iris Mabry, MD, MPH
Agency for Healthcare, Research, and Quality

Tyl Hewitt, PhD
NICHD, NIH

James Panagis, MD, MPH
National Institute of Arthritis and Musculoskeletal and Skin Diseases, NIH

Dennis Wen, MD
Department of Family and Community Medicine
University of Missouri

Appendix 7. Literature Concordance with the Canadian Task Force on Preventive Health Care Review

CTFPHC Force Citation	Inclusion/ Exclusion	Reason for Exclusion
Andersson, 1995	Include	
Barlow, 1962	Include	
Bialik, 1999	Include	
Boere-Boonekamp, 1998	Include	
Burger, 1990	Include	
Clarke, 1989	Include	
Dias, 1993	Include	
Dunn, 1985	Include	
Gardiner, 1990	Include	
Godward, 1998	Include	
Graf, 1984	Include	
Krikler, 1992	Include	
Lennox, 1993	Include	
Macnicol, 1990	Include	
Marks, 1994	Include	
Place, 1978	Include	
Rosendahl, 1995	Include	
Rosendahl, 1994	Include	
Tonnis, 1990	Include	
Tredwell, 1981	Include	
Wedge, 1979	Include	
Berman, 1986	Exclude	Pre-1996 study of screening modalities
Castelein, 1988	Exclude	Pre-1996 study of screening modalities
Catterall, 1994	Exclude	Expert opinion
Cheng, 1994	Exclude	Pre-1996, no comparison of screening modalities
Clarke, 1992	Exclude	Editorial/expert opinion
Davids, 1995	Exclude	Non-primary care based
Dodenhoff, 1996	Exclude	Letter
Donaldson, 1994	Exclude	Commentary
Fulton, 1984	Exclude	Economic Analysis based on data no longer relevant to management
Garvey, 1992	Exclude	Pre-1996 study of screening modalities
Gerscovich, 1997	Exclude	Nonsystematic review

Hansson, 1997	Exclude	Commentary
Harcke, 1993	Exclude	Commentary
Hernandez, 1995	Exclude	Editorial/expert opinion
Hernandez, 1994	Exclude	Poor quality: high risk of bias not accounted for in analysis
Holen, 1994	Exclude	Pre-1996 screening modalities study; all data from study captured and included in Holen 2002
Jomha, 1995	Exclude	Pre-1996, retrospective, nonconsecutive sample in a study of screening modality
Jones, 1989	Exclude	Pre-1996, no comparison of screening modalities
Jones, 1990	Exclude	Pre-1996 study of screening modalities
Langkamer, 1991	Exclude	Case report
Lehmann, 1981	Exclude	Pre-1996, retrospective, no comparison between screening modalities
Mooney, 1995	Exclude	Nonsystematic review
Poul, 1992	Exclude	Pre-1996, no comparison between screening modalities
Poul, 1998	Exclude	No comparison between screening modalities
Rembold, 1998	Exclude	Not specific to DDH
Rosendahl, 1992	Exclude	Pre-1996 study of screening modalities
Secretaries of State for Social Services/ Wales	Exclude	Nonsystematic review; Standing Medical Advisory Committee
Walker, 1977	Exclude	Poor quality: design with high risk of ascertainment bias; reports relative lack of disability but does not quantify
Wolfe, 1990	Exclude	Not specific to DDH
Von Rosen, 1956	Exclude	No functional outcomes
Weinstein, 1987	Exclude	Nonsystematic review
Zieger, 1986a	Exclude	Nonsystematic review, out of scope
Zieger, 1986b	Exclude	Not specific to DDH, no age of patients given
Zieger, 1987	Exclude	Pre-1996 screening, no age of patients given