University of Calgary

FAMILY MEDICINE RESIDENCY
AMBULATORY CLINIC EXPERIENCES

in

• Diagnostic Imaging
• Image-guided Pain Management
• Maternal Fetal Medicine
EFW Locations for ACE’s

Phone (all sites): 403-541-1200

GENERAL DIAGNOSTIC IMAGING (DI)

Primary Site:
*Advanced Medical Imaging Centre (CAMBRIAN)*
2000 Veteran’s Place NW

Alternate Site:
*Beddington Clinic*
#200, 8120 Beddington Blvd NE

IMAGE-GUIDED PAIN MANAGEMENT/MSK (ASCC)

*Advanced Spinal Care Clinic (ASCC at CAMBRIAN)*
located on the 2nd floor at the Advanced Medical Imaging Centre
2000 Veteran’s Place NW

MATERNAL FETAL MEDICINE (MFM)

Primary Site:
*Quarry Park MFM*
#130, 109 Quarry Park Blvd SE

Alternate Site:
*Beddington MFM*
#200, 8120 Beddington Blvd NE

Arrival/End Times for ACE’s

Morning Sessions: 0800h—1200h
Afternoon Sessions: 1300h—1600h
Table of Contents

GENERAL DI
Pre-ACE Self-Evaluation 2
Radiation Exposure 3-4
Plain Radiographs 5
Breast Imaging 6-8
  Breast Article-1 9-20
  Breast Article-2 21-31
General Ultrasound 32-33
Nuclear Medicine 34-36
Post-ACE Self-Evaluation 37
CanMEDS Evaluation 38
Evaluation of General DI ACE 39

IMAGE-GUIDED PAIN MANAGEMENT/MSK (ASCC)
Pre-ACE Self-Evaluation 41
Image-guided Pain Management 42
Post-ACE Self-Evaluation 43
CanMEDS Evaluation 44
Evaluation of ASCC ACE 45
ASCC Article-1 46-58
ASCC Article-2 59-70
ASCC Article-3 71-79

MATERNAL FETAL MEDICINE (MFM)
Pre-ACE Self-Evaluation 81
MFM Questions 82-83
Post-ACE Self-Evaluation 84
CanMEDS Evaluation 85
Evaluation of MFM ACE 86
MFM Article-1 87-101
MFM Article-2 102-122
General Diagnostic Imaging ACE
1. How would you rate your overall knowledge of General DI in the community? (i.e. Types of studies? Indications? Limitations?)

   1  2  3  4  5  6  7  8  9  10
   No knowledge

2. What are 3 goals you would like to achieve during the General DI ACE?
   a) 
   b) 
   c) 

3. What are 3 specific questions you would like to have answered during the General DI ACE?
   a) 
   b) 
   c)
RADIATION EXPOSURE

Background Information:

“Radiation dose” (commonly referred to as “effective radiation dose”):

• used when referring to the radiation risk averaged over the entire body (different tissues have varying sensitivities to radiation exposure)

• Scientific unit of measurement for effective radiation dose: Millisievert (mSv)

“Naturally occurring ‘background’ radiation”:

• What are some examples of “naturally occurring ‘background’ radiation exposure”?

• Average person’s annual effective dose of “naturally occurring ‘background’ radiation? 3 mSv

To put it into perspective:

• Airline flight (~10 hour flight) = 0.03 mSv

• 1 chest x-ray = 0.1 mSv = 3% (~10 days) of annual natural “background” radiation dose

• 1 chest x-ray = 3 airline flights

• Annual natural “background” radiation dose = 30 chest x-rays = 90 airline flights
### Comparison of X-ray and CT Studies with the Effective Dose*

<table>
<thead>
<tr>
<th>Diagnostic Imaging Study</th>
<th>Approximate effective dose</th>
<th>Comparable to # of chest x-rays:</th>
<th>Comparable to natural background radiation for:</th>
<th>Additional lifetime risk from a fatal cancer from the exam**:</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CHEST</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chest x-ray</td>
<td>0.1 mSv</td>
<td>1</td>
<td>10 days</td>
<td>Minimal</td>
</tr>
<tr>
<td>CT chest (routine)</td>
<td>7 mSv</td>
<td>70</td>
<td>2 years</td>
<td>Low</td>
</tr>
<tr>
<td>CT chest (low dose)</td>
<td>1.5 mSv</td>
<td>15</td>
<td>6 months</td>
<td>Very low</td>
</tr>
<tr>
<td><strong>ABDOMEN</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CT abd/pelvis with contrast</td>
<td>10 mSv</td>
<td>100</td>
<td>3 years</td>
<td>Low</td>
</tr>
<tr>
<td>CT colonography (at FMC)</td>
<td>3 mSv</td>
<td>30</td>
<td>1 year</td>
<td>Low</td>
</tr>
<tr>
<td>Abdomen x-ray</td>
<td>6 mSv</td>
<td>60</td>
<td>2 years</td>
<td>Low</td>
</tr>
<tr>
<td><strong>BONE</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spine x-ray</td>
<td>1.5 mSv</td>
<td>15</td>
<td>6 months</td>
<td>Very low</td>
</tr>
<tr>
<td>Extremity x-ray</td>
<td>0.001 mSv</td>
<td>0.1</td>
<td>3 hours</td>
<td>Negligible</td>
</tr>
<tr>
<td><strong>NEURO</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CT head</td>
<td>2 mSv</td>
<td>20</td>
<td>8 months</td>
<td>Very low</td>
</tr>
<tr>
<td>CT spine (entire spine)</td>
<td>6 mSv</td>
<td>60</td>
<td>2 years</td>
<td>Low</td>
</tr>
<tr>
<td><strong>HEART</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cardiac CT Angiography (CTA)</td>
<td>12 mSv</td>
<td>120</td>
<td>4 years</td>
<td>Low</td>
</tr>
<tr>
<td>Cardiac CT Calcium Scoring</td>
<td>3 mSv</td>
<td>30</td>
<td>1 year</td>
<td>Low</td>
</tr>
<tr>
<td><strong>NUCLEAR MEDICINE</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bone Densitometry</td>
<td>0.001 mSv</td>
<td>0.1</td>
<td>3 hours</td>
<td>Negligible</td>
</tr>
<tr>
<td><strong>BREAST</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mammography</td>
<td>0.4 mSv</td>
<td>4</td>
<td>7 weeks</td>
<td>Very low</td>
</tr>
</tbody>
</table>

**Approximate additional risk of fatal cancer for an adult from examination:**

<table>
<thead>
<tr>
<th>Risk Level</th>
<th>**Approximate additional risk of fatal cancer for an adult from examination:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negligible</td>
<td>&lt; 1 in 1,000,000</td>
</tr>
<tr>
<td>Minimal</td>
<td>1 in 1,000,000 — 1 in 100,000</td>
</tr>
<tr>
<td>Very Low</td>
<td>1 in 100,000 — 1 in 10,000</td>
</tr>
<tr>
<td>Low</td>
<td>1 in 10,000 — 1 in 1,000</td>
</tr>
<tr>
<td>Moderate</td>
<td>1 in 1,000 — 1 in 500</td>
</tr>
</tbody>
</table>

**Note:** These risk levels represent very small additions to the 1 in 5 chance we all have of dying from cancer.

**Recommended Reading:**

# PLAIN RADIOGRAPHS

**What are some INDICATIONS for each study?**

<table>
<thead>
<tr>
<th>Study</th>
<th>Indications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chest X-ray</td>
<td></td>
</tr>
<tr>
<td>Abdomen X-ray</td>
<td></td>
</tr>
<tr>
<td>MSK X-rays</td>
<td></td>
</tr>
<tr>
<td>Spine X-rays</td>
<td></td>
</tr>
</tbody>
</table>

**What are some LIMITATIONS for each study?**

<table>
<thead>
<tr>
<th>Study</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chest X-ray</td>
<td></td>
</tr>
<tr>
<td>Abdomen X-ray</td>
<td></td>
</tr>
<tr>
<td>MSK X-rays</td>
<td></td>
</tr>
<tr>
<td>Spine X-rays</td>
<td></td>
</tr>
</tbody>
</table>

*Recommended Reading:*


BREAST IMAGING

Breast Imaging Modalities:

1. What are the three main imaging modalities for breast imaging?

2. What is the primary imaging modality for breast cancer screening?

3. **Mammography:** What kind of things can be better detected on mammography than US?

4. **Ultrasound:** What are some common indications for breast US?

5. **Breast MRI:** What are some common indications for breast MRI?
**Cases: Which diagnostic imaging exam(s) would you request for each scenario and why?**

1. 28 yo F with a palpable “lump” at 12:00 in the left breast. No known risk factors for breast cancer.

2. 55 yo F with a palpable “lump” at 12:00 in the left breast. No strong FHx of breast cancer.

3. 35 yo F with intermittent painful breasts (outer half of both breasts) for “years”.

4. 43 yo F with intermittent painful R breast (upper outer quadrant of right breast) for a few months.

5. 65 yo F with new inversion of left nipple. No other concerns. No FHx of breast cancer.

6. 58 yo F with new yellowish nipple discharge on the right that has occurred twice in the past month. First cousin with breast cancer.

7. 30 yo female with strong FHx of breast cancer (sister with breast cancer at age 37 yo) and BRCA-1/2 positive. Would you order any imaging at this time?

8. 30 yo F with “dense breasts”, but no focal palpable “lumps” or pain.

9. 50 yo F with “dense breasts”, but no focal palpable “lumps” or pain.

10. 31 yo F with “fibrocystic breasts”, but no focal palpable “lumps” or pain and no strong FHx of breast cancer.
11. 48 yo F with “fibrocystic breasts”, but no focal palpable “lumps” or pain and no strong FHx of breast cancer.

12. 14 yo M with new, painless “lump” in retroareolar region on the left.

13. 50 yo M with new, painless “lump” in retroareolar region on the right.

14. 53 yo M with new nipple inversion on the right. Questionable “lump” at 9:00 on the right.

**Recommended Reading:**


The impact of mammographic screening on breast cancer mortality in Europe: a review of observational studies

Mireille Broeders, Sue Moss, Lennarth Nyström, Sisse Njor, Håkan Jonsson, Ellen Paap, Nathalie Massat, Stephen Duffy, Elsebeth Lyne and Eugenio Paci, for the EUROSCREEN Working Group

Objectives To assess the impact of population-based mammographic screening on breast cancer mortality in Europe, considering different methodologies and limitations of the data.

Methods We conducted a systematic literature review of European trend studies (n = 17), incidence-based mortality (IBM) studies (n = 20) and case-control (CC) studies (n = 8). Estimates of the reduction in breast cancer mortality for women invited versus not invited and/or for women screened versus not screened were obtained. The results of IBM studies and CC studies were each pooled using a random effects meta-analysis.

Results Twelve of the 17 trend studies quantified the impact of population-based screening on breast cancer mortality. The estimated breast cancer mortality reductions ranged from 1% to 9% per year in studies reporting an annual percentage change, and from 28% to 36% in those comparing post- and prescreening periods. In the IBM studies, the pooled mortality reduction was 25% (relative risk [RR] 0.75, 95% confidence interval [CI] 0.69–0.81) among invited women and 38% (RR 0.62, 95% CI 0.56–0.69) among those actually screened. The corresponding pooled estimates from the CC studies were 31% (odds ratio [OR] 0.69, 95% CI 0.57–0.83), and 48% (OR 0.52, 95% CI 0.42–0.65) adjusted for self-selection.

Conclusions Valid observational designs are those where sufficient longitudinal individual data are available, directly linking a woman’s screening history to her cause of death. From such studies, the best ‘European’ estimate of breast cancer mortality reduction is 25–31% for women invited for screening, and 38–48% for women actually screened. Much of the current controversy on breast cancer screening is due to the use of inappropriate methodological approaches that are unable to capture the true effect of mammographic screening.

INTRODUCTION

Mammographic breast cancer screening has been the subject of controversy, despite or perhaps due to the fact that it is one of the most scrutinized public health interventions. Randomized controlled trials (RCTs), conducted in the 1970s and 1980s, have shown that mammographic screening can reduce breast cancer mortality by 25–30% after 7–12 years from entry in the trials. Nevertheless, since 2000, concerns have been raised about the validity of these trials because of supposed ‘flaws’ in randomization and ascertainment of cause of death, although these issues have been addressed. More recently, observational studies reporting on the impact of population-based screening programmes have also been questioned. The debate that followed, in academic journals as well as the lay press, has not helped women and their physicians to have a clear view of the benefits of mammographic screening. Concern has also been expressed that women are not fully informed about the potential harms of screening, in particular, possible over-diagnosis of cancers that might not have been diagnosed clinically.

Many countries implemented population-based screening following the results of the RCTs. There are several reasons why the effectiveness of population-based service screening mammography may differ from that observed in the RCTs, including the wider base of professionals who are involved in screening and the improvement of mammographic and other techniques since the trials were conducted. In RCTs and in some observational designs the effect of screening is measured by comparing women invited with women not invited. This comparison is influenced by the attendance rate and therefore reflects the performance of the programme, rather than the screening test itself. The effect estimate will be larger when comparing breast cancer mortality in screened women with that in non-screened women. Service screening effectiveness will also be influenced by the extent of opportunistic screening. Although data on opportunistic screening are scarce, the increased use of
mammography outside organized screening programmes may contribute to a reduction in breast cancer mortality. The emphasis for evaluation has now shifted to population-based screening services, and observational studies will become the main contributors of new information on the impact of breast cancer screening as a public health policy. In this review, we focus on the reduction in breast cancer mortality as the principal benefit of screening, which is by definition a long-term commitment. Several studies corroborate that well-designed observational studies produce results that are similar to those from RCTs. There are, however, specific difficulties in determining the impact of breast cancer screening.

A common first step in the evaluation of screening is to study trends in breast cancer mortality over time. However, the impact of service screening on breast cancer mortality observed in routine population statistics will take many years to emerge. Firstly, with improved treatment, breast cancer survival is generally much higher than in the past while breast cancer incidence has increased in most countries. In combination, the number of deaths in the short-term will be lower, but in the long-term the absolute number of potentially preventable breast cancer deaths has increased. Secondly, it usually takes a number of years before a screening programme is fully implemented. Thirdly, most trend studies are not able to allow for breast cancers diagnosed in women before the start of the screening programme. Finally, when there is no individual data, no corrections can be made for the varying participation behaviour of women invited. Potential confounding, where factors other than screening may also contribute to changes in breast cancer mortality, also presents a complication. Therefore, service-based screening programmes cannot be evaluated using only analyses of trends.

A further difficulty in determining the impact of screening is the typical absence of a readily available control population. Studies which were able to identify, albeit for a limited time period, a group of contemporaneous controls that were not (yet) invited for screening have mostly used the incidence-based mortality (IBM) approach. IBM studies estimate the impact of screening by calculating mortality rates based on breast cancer deaths occurring in women with breast cancer diagnosed after their first invitation to screening. Using individual data in IBM studies can overcome many of the problems that affect trend analyses.

Case-control (CC), or case-referent, studies have also been used to evaluate the impact of service screening. A CC study compares breast cancer deaths (cases) with a sample of women who have not died from breast cancer, in terms of individual screening exposure. There is an efficiency gain in taking a sample of the population invited to be screened, rather than observing the entire population. If correctly designed and analysed, the CC approach offers a valid and efficient method for estimating the impact of service screening programmes.

Our objective is to assess the impact of population-based screening with mammography on breast cancer mortality in Europe. A best estimate for the effectiveness of population-based screening in Europe will be provided, acknowledging the different methodologies and the limitations of the available data.

METHODS

A systematic search of PubMed was performed based on all papers published up to February 2011 (details in the Appendix A). We identified 5011 English-language articles evaluating the effect of mammographic screening on breast cancer mortality in Europe. After inspection of titles and abstracts, 122 studies were considered to be relevant. These were reviewed and further selected using the following criteria: (a) the study represents original data on a population-based screening programme in Europe, (b) breast cancer mortality is reported, (c) the analysis includes at least some of the age groups between 50 and 69, and (d) one of the following observational research designs was used: trend, IBM or CC study. In addition, we only considered studies estimating the impact of current breast cancer screening programmes, and therefore excluded those which had less than three years’ overlap with the relevant current regional or national population screening programme. Based on these criteria, 83 studies were excluded on the following grounds: data from RCTs (n = 17), outcome measure is not breast cancer mortality (n = 20), insufficient overlap with current population-based programme (n = 11), data limited to younger or older women (n = 9), study reporting no new data or no analysis with regard to screening (n = 15), modelling study (n = 6), full paper not in English (n = 2), study on opportunistic screening (n = 2) and study on benign breast disease (n = 1).

In addition to the literature search, the Working Group added publications fulfilling the inclusion criteria but not identified by the search and new publications that became available after February 2011 (n = 5). Studies were summarized according to the three designs (see Table 1): trend studies, IBM studies and CC studies.

Trend studies

Relevant papers were those that reported on trends in breast cancer mortality rates in a population as a whole in relation to the introduction and/or extent of population based mammographic screening (n = 17). They are described in detail elsewhere in this supplement. These studies were usually based on aggregated data obtained from routine sources, such as cancer registries. Trend studies were either classified into (a) descriptions of the trend over time in breast cancer mortality in relation to the timing of the introduction of population-based screening (n = 5), or (b) those which included a more detailed analysis with the aim of quantifying the impact of screening on mortality (n = 12). Methods of analysis in the latter category included Poisson regression (with or without age cohort modelling), and the use of joint-point regression to identify ‘break points’ at which changes in mortality trends occurred (see Table 2). Due to the varied methodology and comparisons in the studies, no attempt was made to produce a pooled estimate of the effect of screening.

IBM studies

In an IBM study all breast cancer deaths occurring in a dynamic or cohort population over a period of time are
Table 1  Publications on the impact of population-based screening with mammography in Europe according to observational study design

<table>
<thead>
<tr>
<th>Country</th>
<th>Start of national programme</th>
<th>Age group targeted (years)</th>
<th>Study design (first author, journal and year of publication)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Antilla, BMC Public Health (2008)</td>
</tr>
<tr>
<td>Italy</td>
<td>1995</td>
<td>50–69</td>
<td>Barchielli, Cancer Causes Control (2001)</td>
</tr>
<tr>
<td>16 European countries</td>
<td></td>
<td></td>
<td>Blanks, BMJ (2000)</td>
</tr>
<tr>
<td>Nordic capitals</td>
<td></td>
<td></td>
<td>Quinn, BMJ (1993)</td>
</tr>
<tr>
<td>30 European countries</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Northern Ireland</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>the Netherlands, Sweden</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1In half of the Swedish counties, the lower age limit is 40; in the other half screening starts at age 50
2Current age limits are 50–70 but will be extended to 47–73
3Northern Ireland (UK) compared with the republic of Ireland, the Netherlands compared with Belgium/Flanders and Sweden compared with Norway

<table>
<thead>
<tr>
<th>Study design (first author, journal and year of publication)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trend studies (n = 17)</td>
</tr>
<tr>
<td>IBM studies (n = 20)</td>
</tr>
<tr>
<td>Case-control studies (n = 8)</td>
</tr>
</tbody>
</table>
Table 2  Summary of European trend studies that report an estimate of the effect of screening

<table>
<thead>
<tr>
<th>Service screening programme</th>
<th>Reference</th>
<th>Study area</th>
<th>Start</th>
<th>100% coverage</th>
<th>Age group invited</th>
<th>Time period studied</th>
<th>Age range studied</th>
<th>Method</th>
<th>Reference group</th>
<th>Reduction in breast cancer mortality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Denmark</td>
<td>Jørgensen (2010)</td>
<td>Copenhagen and Funen</td>
<td>1991, 1993</td>
<td>50–69</td>
<td>1971–2006</td>
<td>35–84</td>
<td>Poisson</td>
<td>Rest of Denmark 55–74 versus other age groups</td>
<td>1977–2006: 1% (95% CI – 1 to 4) reduction per annum versus 2% (95% CI 1–3) in non-screening areas</td>
<td></td>
</tr>
<tr>
<td>Spain</td>
<td>Calanes (2009)</td>
<td>Spain</td>
<td>1990–99</td>
<td>2001</td>
<td>(45)50–64</td>
<td>1980–2006</td>
<td>25+</td>
<td>Joinpoint</td>
<td>Age group</td>
<td>Changepoint 1993. Reduction pa (95% CI) 1993–2006 by age group: 4% (95% CI: 3.3–4.4) pa 25–44 years; 3.1% (95% CI: 2.9–3.4) pa 45–64 years, 1.3% 65+ years</td>
</tr>
<tr>
<td></td>
<td>Pons-Vigués (2008)</td>
<td>Barcelona Spain</td>
<td>1995</td>
<td>2004</td>
<td>50–69</td>
<td>1984–2004</td>
<td>50–74</td>
<td>Poisson</td>
<td>Before/after; grouped by start date.</td>
<td>1995–2004: 5% (95% CI 1–8) reduction pa versus 1% (95% CI 1–2%) before start</td>
</tr>
</tbody>
</table>

CI, confidence interval

*Countries (UK) compared with the republic of Ireland, Sweden compared with Norway and the Netherlands compared with Belgium/Flanders*
enrolled in the study only if the breast cancer diagnosis occurred in a certain time/age window (taking into account eligibility and opportunity to be screened) and the population is classified by screening or by invitation to screening. Thus, for example, breast cancer deaths in the 15 years after screening is initiated in one region, from tumours diagnosed in that 15-year period, may be compared with the corresponding deaths from tumours diagnosed in the same period in a region without screening. The selection of IBM studies contributing to this overall review is described in detail elsewhere in this supplement.23 There were 20 IBM studies – one each from Denmark, Norway and Spain, two from Italy, seven from Finland and eight from Sweden. A key issue in these studies is how the breast cancer mortality expected in the absence of screening is estimated. Another methodological concern is how the study deals with potential biases in the estimated mortality reduction due to screening. Because breast cancer cases are diagnosed earlier in screened women than in those who are not screened, a longer follow-up period for breast cancer deaths than the accrual period for cases will confer an artificial increase in mortality in the screening period due to fatal cases whose diagnosis is moved to the accrual period due to lead time. The same consideration applies to age at diagnosis. If mortality includes deaths from tumours diagnosed within a certain age range, but with no upper limit on age at death, there will be a number of fatal cancers diagnosed by screening within the age range, which would otherwise have been excluded as diagnosed symptomatically above the age range.23

Table 3 presents some basic characteristics of the IBM studies. Where there was overlapping data, the study used in this review was selected on the basis of follow-up time, judgement of quality of the comparison group and study size. We calculated a pooled estimate of the effect on breast cancer mortality in women invited versus not invited, as well as a pooled estimate for women screened versus not screened, using the formula described by Duffy et al.72 The effect sizes were pooled using the inverse variance method (random effects model) and heterogeneity between the studies was assessed.14,73

**CC studies**

A CC study is embedded in a cohort or a dynamic population and based on sampling of the population experience. Breast cancer deaths (cases) in the population are collected over the period of interest and controls who have not died of breast cancer are selected from the same population, often closely matched by temporal factors. Breast cancer cases and control subjects are then compared with respect to screening history before the date of diagnosis of the breast cancer case. The eight CC studies used in this review (Table 4) came from a recently published methodological overview, but we excluded non-European studies26 and added publications by Broeders et al.50 and Otto et al.47

The results were pooled to obtain estimates of the effect on breast cancer mortality for women screened versus not screened, based on the crude odds ratios (ORs) as well as ORs adjusted for self-selection. In addition, intention to treat estimates were calculated, using the formula described by Duffy et al.72 in order to compare the women invited with those not invited. Because the studies by Broeders et al. and van Schoor et al. were both conducted in Nijmegen, with overlap in the included cases, the former was excluded from the meta-analysis. The effect sizes were pooled as above.14,73

**Breast cancer mortality as an outcome measure**

Breast cancer mortality is the most appropriate primary end-point for evaluating screening, although its use has been questioned.74,77 An outcome parameter which avoids problems with cause of death classification is (refined) excess mortality from breast cancer, which includes all mortality associated with breast cancer, even indirectly caused deaths, such as treatment-induced mortality, or deaths caused by the stress imposed by the cancer.78 However, this method, so far, has only been used in Sweden.

Potential limitations of using breast cancer mortality as an outcome measure are that there could be an increase in deaths attributed to breast cancer because more breast cancer cases are diagnosed in screened women, and the misclassification of breast cancer as the underlying cause of death because the treating physician is influenced by the screening history of the patient. Screening may also affect mortality from other causes, for example, due to complications arising from procedures triggered by screening.79 However, several studies explicitly assessed the quality of cause-of-death determination in relation to mammographic screening and found no significant evidence of bias.77–80

**RESULTS**

**Trend studies**

Of the 12 trend studies, three used joinpoint regression, and nine Poisson regression (Table 2). Five papers were based on all of an individual country (England, the Netherlands and Spain), two studied the programme in the city of Florence (Italy), two studied different regions in Spain and one studied two regions of Denmark. One paper included Northern Ireland, the Netherlands and Sweden in comparison with the Republic of Ireland, Belgium/Flanders and Norway, respectively. The most recent paper studied nine counties in Sweden.

Authors of several studies estimated the annual percentage change in mortality, while others presented a comparison between two distinct time periods. Of the former, estimates ranged from reductions of 1% to 9% per year; for those studies with adequate follow-up (at least 10 years from the date of full coverage by invitation) the estimates were 1%, 2.3–2.8% and 9%.31,46,48,52–55 Of the three studies comparing time periods within a single country, all had adequate follow-up, and the estimates of mortality reduction compared with a prescreening period ranged from 28% to 36%.41,53,64

**IBM studies**

Table 3 shows the design characteristics of the IBM studies. The outcomes were generally compatible when differences...
Table 3  Design characteristics of European IBM studies, excluding those with overlapping data, and estimate of effect

<table>
<thead>
<tr>
<th>Reference</th>
<th>Region and age group</th>
<th>Maximum follow-up (years)</th>
<th>Accrual period = follow-up</th>
<th>Age at diagnosis = age at breast cancer death</th>
<th>Contemporaneous uninvited comparison group</th>
<th>Balanced follow-up Comparison</th>
<th>Lead time adjustment</th>
<th>Relative risk, invited to screening (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Denmark</td>
<td>Olsen (2005)³²</td>
<td>Copenhagen, 50–69</td>
<td>11</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>NR</td>
<td>0.75 (0.63–0.89)</td>
</tr>
<tr>
<td>Finland</td>
<td>Sarkeala, Br J Cancer (2008)³⁶</td>
<td>8 municipalities, 50–69</td>
<td>12</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>NR</td>
<td>0.72 (0.51–0.97)</td>
</tr>
<tr>
<td></td>
<td>Hakama (1997)³⁹</td>
<td>2/3 of municipalities, 50–64</td>
<td>6</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>NR</td>
<td>0.76 (0.53–1.09)</td>
</tr>
<tr>
<td>Italy</td>
<td>Paci, Eur J Cancer (2002)⁴⁵</td>
<td>Florence, 50–69</td>
<td>10</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>NA</td>
<td>0.81 (0.64–1.01)</td>
</tr>
<tr>
<td>Norway</td>
<td>Kalager (2010)⁵¹</td>
<td>4 counties, 50–69</td>
<td>10</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>0.88 (0.73–1.05)</td>
</tr>
<tr>
<td>Spain</td>
<td>Ascunce (2007)⁵³</td>
<td>Navarra, 45–65</td>
<td>11</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>0.58 (0.44–0.75)</td>
</tr>
<tr>
<td>Sweden</td>
<td>SOSSEG (2006)⁵⁹</td>
<td>13 areas, 50–69</td>
<td>22</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>0.73 (0.69–0.77)</td>
</tr>
</tbody>
</table>

NR, not required; NA, not adjusted; IBM, incidence-based mortality; CI, confidence interval
<table>
<thead>
<tr>
<th>Study</th>
<th>Geographical area</th>
<th>Study population</th>
<th>Age group</th>
<th>Time period for including cases (breast cancer deaths)</th>
<th>Screening exposure</th>
<th>Crude odds ratio (95% CI)</th>
<th>Adjusted estimates (95% CI)</th>
<th>Using own correction factor</th>
<th>SES</th>
<th>Participation rate</th>
<th>Intention to treat estimate (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Iceland</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gabe (2007)**</td>
<td>Iceland</td>
<td>226/902</td>
<td>≥40</td>
<td>1990–2002</td>
<td>Ever/never</td>
<td>0.59 (0.41–0.84)</td>
<td>0.65 (0.39–1.09)</td>
<td>(factor = 1.17, 95% CI 1.08–1.26) — also adjusted for screening-opportunity bias</td>
<td>62%</td>
<td>0.87</td>
<td>(0.72–1.06)</td>
</tr>
<tr>
<td><strong>Italy</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Puliti (2008)**</td>
<td>Five regions</td>
<td>657/2058</td>
<td>50–74</td>
<td>1988–2002</td>
<td>Ever/never</td>
<td>0.46 (0.38–0.56)</td>
<td>0.55 (0.36–0.85)</td>
<td>(factor = 1.11, 95% CI 0.87–1.40)</td>
<td>65%</td>
<td>0.72</td>
<td>(0.56–0.93)</td>
</tr>
<tr>
<td><strong>The Netherlands</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Otto (2012)**</td>
<td>Southwest region</td>
<td>755/3739</td>
<td>49–75</td>
<td>1990–2003</td>
<td>Ever/never</td>
<td>0.45 (0.37–0.54)</td>
<td>0.51 (0.40–0.66)</td>
<td>(factor = 1.11, 95% CI 0.99–1.25)</td>
<td>75%</td>
<td>0.65</td>
<td>(0.56–0.77)</td>
</tr>
<tr>
<td>Van Schoor (2011)**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Ever/never in index invitation + previous invitation (4 years, maximum 2 screens)</td>
<td>0.35 (0.19–0.64)</td>
<td>0.28 (0.12–0.60)</td>
<td>(factor = 0.84, 95% CI 0.58–1.21)</td>
<td>from Paap et al.**</td>
<td>68%</td>
<td>0.47</td>
</tr>
<tr>
<td>Paap (2010)**</td>
<td>Midsouth Limburg</td>
<td>118/118</td>
<td>50–75</td>
<td>2004–2005</td>
<td>Index invitation</td>
<td>0.30 (0.14–0.63)</td>
<td>0.24 (0.10–0.58)</td>
<td>(factor = 0.84, 95% CI 0.58–1.21)</td>
<td>82%</td>
<td>0.36</td>
<td>(0.20–0.64)</td>
</tr>
<tr>
<td>Broeders (2002)**</td>
<td>Nijmegen</td>
<td>26/192, 36/231, 48/194, 34/118, 12/39</td>
<td>40–49, 50–59, 60–69, 70–79, &gt;79</td>
<td>1987–1997</td>
<td>Ever/never in index invitation + previous invitation (4 years, maximum 2 screens)</td>
<td>0.90 (0.38–2.14), 0.71 (0.35–1.46), 0.60 (0.42–1.54), 1.13 (0.50–2.58), 2.92 (0.55–15.40)</td>
<td>0.24 (0.10–0.58)</td>
<td>(factor = 0.84, 95% CI 0.58–1.21)</td>
<td>59%</td>
<td>0.85</td>
<td>(0.63–1.15)**</td>
</tr>
<tr>
<td><strong>United Kingdom</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Allgood (2008)**</td>
<td>East Anglia</td>
<td>284/568</td>
<td>50–70</td>
<td>1995–2004</td>
<td>Ever/never</td>
<td>0.35 (0.24–0.50)</td>
<td>0.35 (0.23–0.51)</td>
<td>(factor = 0.85, 95% CI 0.63–1.15)</td>
<td>80%</td>
<td>0.65</td>
<td>(0.48–0.88)</td>
</tr>
<tr>
<td>Fielder (2004)**</td>
<td>Wales</td>
<td>419/717</td>
<td>50–75</td>
<td>1998–2001</td>
<td>Ever/never</td>
<td>0.62 (0.47–0.82)</td>
<td>0.75 (0.49–1.14)**</td>
<td>No difference</td>
<td>77%</td>
<td>0.96</td>
<td>(0.73–1.27)</td>
</tr>
</tbody>
</table>

**Based on original publication, except for van Schoor (personal communication) and Broeders (personal communication)**

†Calculated using the formula by Duffy et al. (Appl Stat 2002) and using the crude OR

††Index invitation is the most recent invitation before diagnosis of the breast cancer case

‡‡Limited to cancers diagnosed in 1995–2001, where the crude OR was 0.49 (0.36–0.66)

†††Based on an overall crude OR of 0.64 (0.44–0.92) (Broeders, personal communication) and self-selection factor of 1.08, 95% CI 0.85–1.37 (Paap et al.**).
in methodology and local circumstances were taken into account. Details are given elsewhere in this supplement. Those with the strongest designs had (a) expected breast cancer mortality estimated from a cohort of women not yet invited or from historical and contemporaneous control groups, and (b) an accrual period equal to the follow-up period for breast cancer deaths. Using all IBM studies, excluding overlapping data-sets, produced a pooled relative risk (RR) estimate of 0.75 (95% confidence interval [CI] 0.69–0.81) for invitation to screening, with no significant heterogeneity (P = 0.23). The combined RR for women actually screened was 0.62 (95% CI 0.56–0.69), again with no significant heterogeneity (P = 0.40). Figure 1 shows the forest plots.

CC studies

Of the eight CC studies included, one came from Iceland, one from Italy, four from the Netherlands and two from the UK (Table 4), but their designs were very similar. The definition of exposure to screening was based on a comparison of women ‘ever’ screened versus women ‘never’ screened in four studies. All Dutch studies adopted the concept of the index invitation, defined as the invitation date closest to the date of diagnosis of the case. The comparison in these studies was between women screened in an exposure period which varied from one to three screening examinations versus women not screened in this period. All studies reported ORs adjusted for self-selection bias, either using the correction factor estimated by Duffy et al. or their own correction factor, all closer to 1 than the Duffy factor. Based on the results in the original publications, we also calculated the reduction in breast cancer mortality for women invited versus not invited.

Seven CC studies were included in a pooled analysis (see Methods). The combined unadjusted OR was 0.46 (95% CI 0.40–0.54), a significant 54% reduction in breast cancer mortality for screened versus not screened women. This became a 48% reduction after adjusting for self-selection (OR 0.52, 95% CI 0.42–0.65). There was no evidence of heterogeneity in either analysis (P = 0.10 and 0.17, respectively). The combined mortality reduction for invitation to screening was 31% (OR 0.69, 95% CI 0.57–0.83), but with significant heterogeneity (P = 0.005). Figure 2 shows the forest plots. The squares representing the point estimates in the individual CC studies are proportional to the precisions of the log ORs. The order of these may vary when adjusted for self-selection bias as after adjustment the precision also depends on the standard error of the self-selection correction. This in turn depends on the participation rate in each study.

DISCUSSION

Our overview indicates that the estimates from observational studies, using different study designs, are consistent with a

Figure 1 Incidence-based mortality studies excluding overlapping data: (a) estimates for breast cancer mortality reduction in women invited versus not invited; (b) estimates for breast cancer mortality reduction in women screened versus not screened; (c) crude odds ratios for breast cancer mortality reduction translated to intention to treat estimates for women invited versus not invited.

Figure 2 Case-control studies excluding overlapping data: (a) crude odds ratios for breast cancer mortality reduction in women screened versus not screened; (b) crude odds ratios for breast cancer mortality reduction, corrected for self-selection, in women screened versus not screened; (c) crude odds ratios for breast cancer mortality reduction translated to intention to treat estimates for women invited versus not invited.
breast cancer mortality reduction of 25–31% for women in Europe invited for population-based screening. The current best estimate of the effectiveness of European screening programmes is therefore at least as large as that observed in the long-term follow-up of the Swedish RCTs\(^{86}\) or more recent meta-analyses.\(^{24,82}\)

Given the methodological limitations inherent in observational studies, and the differences in designs, the similarity in the effect estimates from trend, IBM and CC studies is noteworthy. Using all IBM studies without overlapping data, the reduction in breast cancer mortality for women invited was 25%. The corresponding intention to treat estimate in the CC studies was 31%. The relative reduction in breast cancer mortality for women who actually participated in screening was 38% based on IBM studies and 48% based on CC studies. Of the three trend studies comparing time periods within a single country, all had adequate follow-up, and the estimates of mortality reduction compared with a prescreening period ranged from 28% to 36%.

The choice of IBM studies to include in the case of overlapping data was not crucial to the estimated mortality reduction, because pooling all studies, including those with overlapping data, gave a mortality reduction of 24%, and selection of three studies on the basis of both historical and contemporaneous comparison groups gave a reduction of 26%.\(^{23}\) The heterogeneity among studies of the intention to treat estimate from the CC studies is likely to be due to differing uptake rates between studies, because there was no significant heterogeneity when the effect of actually being screened was assessed.

The study and analysis of population breast cancer mortality rates can be a first step in evaluating the impact of screening on mortality. However, such analyses should be restricted to the age ranges likely to demonstrate a benefit from screening; they should attempt to exclude time periods where dilution due to deaths in women diagnosed preinvitation will be evident; and they should attempt to take account of past underlying trends. We do not support the recommendation of Harris et al.\(^{22}\) to focus on a trend or ecological approach.

The most valid observational designs are those where longitudinal individual data are available, directly linking screening history to the cause of death, achieved using either an IBM or a CC approach. IBM studies and CC studies have one major feature in common – they typically take as clinical endpoint deaths from cancers which have been diagnosed in the age range and time period in which screening is offered. This avoids dilution bias associated with deaths from breast cancers in a given period from tumours diagnosed before that period began.\(^{62}\) The most obvious difference between the two is that the CC study is retrospective and the IBM study prospective.

In the CC study, data on deaths from the cancer in question are collected along with that from subjects who have not died of the disease, and screening histories retrieved retrospectively. There are a number of well-known potential biases associated with this design, some conservative and some anticonservative.\(^{24,84}\) However, these can be minimized by appropriate design or corrected for in the statistical analysis.\(^{25,85}\) Some biases, such as residual confounding after adjusting for age, tend to be very small.\(^{86}\)

Typically in the IBM studies, rates of death from cancers diagnosed in a population and period of invitation to screening are compared with the corresponding rates in a population or period without such invitation.\(^{39}\) This too has potential biases. There is likely to be confounding of some variables between populations and periods if individual data on invitation and screening are not available. For example, if a before-after comparison of IBM is carried out, the time cut-off will inevitably incur some misclassification of exposure to invitation, because screening is usually phased in over a period of years.\(^{86}\) In the CC approach, individual screening histories are retrieved so there is no misclassification of exposure.\(^{25}\)

In principle, screening exposure can be ascertained for all subjects in the population in the IBM approach, but this involves retrieval of data on tens or even hundreds of thousands of subjects, whereas the CC design typically involves much smaller numbers.\(^{15}\) Therefore, the CC approach is a more economic research strategy, even though it may involve more complex design or analytic procedures. However, if exposure to screening is ascertained for all study subjects on an individual basis in both study designs, the intention-to-treat estimate from CC studies should be similar to that from the IBM studies, as indeed is observed in this review.

**CONCLUSION**

After considering all published data from European studies, the reduction in breast cancer mortality associated with mammographic population-based service screening programmes is in the range of 25–31% for women invited for screening and 38–48% for women actually screened with sufficient follow-up time. It appears that much of the current controversy surrounding the value of mammographic screening is due to the use of inappropriate methodological approaches that are unable to capture the true effect of mammographic screening.

**EUROSCREEN WORKING GROUP**

**Members:**
- Ancelle-Park R (F)\(^1\), Armaroli P (I)\(^2\), Ascunce N (E)\(^3\), Bisanti, L (I)\(^4\), Bellisario C (I)\(^5\), Broeders M (NL)\(^6\), Cogo C (I)\(^6\), De Koning H (NL)\(^7\), Duffy S W (UK)\(^8\), Frigerio A (I)\(^9\), Giordano L (I)\(^2\), Hofvind S (N)\(^9\), Jonsson H (S)\(^10\), Lynge E (DK)\(^11\), Massat N (UK)\(^9\), Miccinesi G (I)\(^12\), Moss S (UK)\(^9\), Naldoni C (I)\(^13\), Njor S (DK)\(^11\), Nystrom L (S)\(^14\), Paap E (NL)\(^2\), Pacì E (I)\(^12\), Patrick J (UK)\(^15\), Ponti A (I)\(^2\), Puliti D (I)\(^12\), Segnan N (I)\(^2\), Von Karsa L (D)\(^16\), Tornberg S (S)\(^17\), Zappa M (I)\(^12\), Zorzi M (I)\(^6\)

**Affiliations:**
- \(^1\)Ministère du travail de l’emploi et de la santé, Paris, France
- \(^2\)CPO-Piedmont, Turin, Italy
- \(^3\)Navarra Breast Cancer Screening Programme, Pamplona, Spain
- \(^4\)S.C. Epidemiologia, ASL di Milano, Italy

Journal of Medical Screening 2012 Volume 19 Suppl 1
The Netherlands to host a meeting of the EUROSCREEN mortality working group in July 2011.

REFERENCES


Journal of Medical Screening 2012 Volume 19 Suppl 1


Anttila T, Sarkeala T, Hakulinen T, Heinävaara S. Impacts of the Finnish service screening programme on breast cancer rates. BMC Public Health 2006;8:38


APPENDIX A

Search Strategy

Evaluation of the effect of service screening programmes with mammography on the breast cancer mortality in Western Europe

SEARCH STRATEGY

We searched the National Library of Medicine PubMed with the following search terms:

1. 'Mortality'[Mesh]
2. 'Mass Screening'[Mesh]
3. 'Mammography'[Mesh]
4. 'Breast Neoplasms/mortality'[Mesh]
5. breast cancer mortality
6. screening
7. mammography
8. ((#1 OR #4) OR #5)
9. (#2) OR #6
10. (#7) OR #3
11. ((#8) AND #9) AND #10

This search strategy retrieved a total of 2462 papers.

12. effect'
13. evaluation
14. impact
15. trend
16. service screening
17. programme screening
18. mass screening
19. breast cancer
20. mortality
21. survival
22. ((#12) OR #13) OR #14) OR #15
23. ((#16) OR #17) OR #18
24. (#20) OR #21
25. (((#22) AND #23) AND #19) AND #24

This search strategy retrieved a total of 1680 papers.

26. 'Mortality/trends'[Mesh]
27. 'Survival Analysis'[Mesh]
28. 'Survival Rate/trends'[Mesh]
29. ((#26) OR #27) OR #28
30. ((#29) AND #2) AND #4

This search strategy retrieved a total of 193 papers.

31. PubMed 'related articles' to the following articles suggested by experts in the field, not retrieved by the previous search strategies:


This search strategy retrieved a total of 726 papers.

These searches were supplemented with suggestion by experts in the field.

The results were sorted by Europe Western Countries: The Netherlands, Finland, Sweden, Norway, Iceland, Denmark, UK, Ireland, Germany, Austria, Italy, Spain, Greece, Nordic countries, Europe (not specified).

We considered all articles published in English language up to February 2011 (no date restriction); the articles were imported into ProCite to select the papers considered relevant after the reading of title and abstracts.
Mammographic Screening and Breast Cancer Mortality: A Case-Control Study and Meta-analysis

Carolyn Nickson, Kate E. Mason, Dallas R. English, et al.


Updated Version
Access the most recent version of this article at:
doi:10.1158/1055-9965.EPI-12-0468

Supplementary Material
Access the most recent supplemental material at:
http://cebp.aacrjournals.org/content/suppl/2012/09/05/21.9.1479.DC1.html

Cited Articles
This article cites 45 articles, 18 of which you can access for free at:
http://cebp.aacrjournals.org/content/21/9/1479.full.html#ref-list-1

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.
Mammographic Screening and Breast Cancer Mortality: A Case–Control Study and Meta-analysis

Carolyn Nickson¹, Kate E. Mason¹, Dallas R. English², and Anne M. Kavanagh¹

Abstract

Background: Observational studies are necessary to assess the impact of population screening on breast cancer mortality. While some ecological studies have notably found little or no association, case–control studies consistently show strong inverse associations, but they are sometimes ignored, perhaps due to theoretical biases arising from the study design. We conducted a case–control study of breast cancer deaths in Western Australia to evaluate the effect of participation in the BreastScreen Australia program, paying particular attention to potential sources of bias, and undertook an updated meta-analysis of case–control studies.

Methods: Our study included 427 cases (women who died from breast cancer), each matched to up to 10 controls. We estimated the association between screening participation and breast cancer mortality, quantifying the effect of potential sources of bias on our findings, including selection bias, information bias, and confounding. We also conducted a meta-analysis of published case–control studies.

Results: The OR for participation in the Western Australian BreastScreen program in relation to death from breast cancer was 0.48 [95% confidence interval (CI), 0.38–0.59; P < 0.001]. We were unable to identify biases that could negate this finding: sensitivity analyses generated ORs from 0.45 to 0.52. Our meta-analysis yielded an OR of 0.51 (95% CI, 0.46–0.55).

Conclusions: Our findings suggest an average 49% reduction in breast cancer mortality for women who are screened. In practice, theoretical biases have little effect on estimates from case–control studies.

Impact: Case–control studies, such as ours, provide robust and consistent evidence that screening benefits women who choose to be screened. Cancer Epidemiol Biomarkers Prev; 21(9); 1479–88. ©2012 AARC.

Introduction

Trials of mammographic screening for breast cancer show that screening reduces breast cancer mortality by 25% (1). With population screening programs now established in at least 27 countries (2), observational study designs can be used to estimate the impact of population screening programs. These designs include case–control studies comparing prior screening participation by individual women who have died from breast cancer and living women, cohort studies comparing breast cancer mortality between screened and unscreened women, and ecological studies relating population levels of breast cancer mortality and screening activity without information on which individuals were screened and which were not (3). Assuming a benefit of screening, case–control and cohort studies would be expected to give stronger associations than ecological studies, except if participation rates are close to 100%.

The Australian population-based screening program BreastScreen Australia has provided free biennial mammographic screening since the mid-1990s. Since 1996, biennial participation in the target age group of 50 to 69 years has ranged from 52% to 57% (4). Four observational studies in Australia have yielded highly varied estimates of the association between screening and breast cancer mortality. A case–control study from South Australia reported an OR of 0.53 [95% confidence interval (CI), 0.40–0.70; ref. 5]. Two ecological studies comparing screening participation rates and breast cancer mortality for small areas within New South Wales (6) and the whole of Australia (7) found that higher participation was associated with lower breast cancer mortality after adjustment for several potential confounding variables. Another more recent ecological study found that the screening program has had little effect on mortality (8) and although an expert committee rejected this design as too crude (7), the study received considerable attention in the media as evidence that mammographic screening programs are ineffective or less effective than previously described (9, 10).

A recently published European ecological study that found no mortality benefit from screening (11) also received considerable attention and reignited the debate...
about the efficacy of screening (12–15). The study and the subsequent media coverage paid little heed to numerous contrary results from randomised trials and previous observational studies (including studies in the countries evaluated; ref. 16, 17), and a subsequent study by the same authors finding screening participation was associated with a 16% reduction in mortality in Sweden (18) maintained a low profile. Meanwhile, case-control studies have consistently demonstrated that participation in screening is associated with lower breast cancer mortality (19, 20) with a previous meta-analysis yielding a pooled odds ratio (OR) of 0.44 (95% CI 0.38–0.50; ref. 20).

To evaluate the effect of participation in the BreastScreen Australia program, we conducted a case–control study of deaths from breast cancer in Western Australia. To evaluate the potential effect of biases discussed in the study of deaths from breast cancer in Western Australia. We also conducted a meta-analysis of case–control studies evaluating mammographic screening.

Materials and Methods

Primary analysis

Participants. In the BreastScreen Australia program, women aged 50 to 69 are directly invited to attend for screening from their registration on the Electoral Roll and then reinvited to their next screen [ref. 21]. Registration to vote is compulsory in Australia, and in 2005, an estimated 98.9% of the eligible population in Western Australia was registered (22). Women aged 40 to 49 years are also eligible to attend but are not directly invited.

For consistency with the program’s target age range, the source population comprised all women on the Western Australia Electoral Roll between 1995 and 2006 who were 50 years of age or older at some time during that period.

To identify cases and potential controls, the Data Linkage Branch of the Department of Health Western Australia used standard protocols (23) to link data from the Electoral Roll to BreastScreen Western Australia screening records and the population-based Western Australia Cancer Registry (which routinely links cancer diagnoses to the national death registry). Cases were women in the source population who died from breast cancer between 1995 and 2006. For each case, incidence density sampling was used to randomly select 10 controls from women in the source population who were alive on the date the case died, resident in Western Australia at the time of the case’s diagnosis, and born in the same month and year as the case (incidence density sampling involves matching each case to a sample of those who are at risk of becoming a case at the time of case occurrence; refs. 24, 25). Controls were not excluded if they had a breast cancer diagnosis. Women were assumed to be alive if the Western Australia Electoral Roll had no record of their death.

For each woman, we collected information on date of birth, screening history, date of any cancer diagnosis and, for cases, date of death. On the basis of place of residence, we assigned to each woman indicators of relative socioeconomic disadvantage (as a proxy for individual socioeconomic status) and remoteness from health services. Each case–control set was assigned a reference date that was the date of the earliest breast cancer diagnosis for that set.

Definition of exposure to screening with BreastScreen. Women were defined as exposed to screening if they received screening mammograms from BreastScreen at any time between their 50th birthday and their case–control set’s reference date. Mammograms after the reference date were ignored to ensure equal opportunity for screening within the case–control set (26, 27). Because BreastScreen is intended for asymptomatic women and to avoid misclassifying diagnostic tests as screening (28), the small number of mammograms for which women presented with symptoms (specifically nipple discharge and/or breast lumps) were ignored.

Statistical analyses. Case–control sets were excluded if their reference date was before 1995 (thus excluding from the analysis women with a cancer diagnosis before potential screening exposure) or if the women were younger than 50 or older than 69 years at the reference date. Women were excluded from the primary analysis if they first attended BreastScreen for screening before 50 years of age. If a case was excluded, so were her matched controls.

All analyses were undertaken using Stata 11.0 (StataCorp). We used conditional logistic regression analysis to estimate the OR for participation in screening in relation to breast cancer mortality. Analyses were adjusted for confounding by socioeconomic status and remoteness from health services using proxy measures based on place of residence of each woman. The data linkage service determined each woman’s place of residence at the small-area level of Census Collector District (CCD; ref. 29) and on the basis of this assigned women a score indicating relative socioeconomic disadvantage (based on the Socio-Economic Index For Areas (SEIFA); ref. 30) and remoteness (based on the Accessibility/Remoteness Index of Australia (ARIA); ref. 31). These scores were categorized into quintiles and used to infer individual socioeconomic status and geographical access to health services (recognizing that the use of proxy measures will result in some misclassification) in the closest Census year before the reference date. For the small number of women (18 cases and 195 controls) for whom data were unavailable at the CCD level, SEIFA and ARIA values were specified using larger spatial units (Statistical Local Areas (SLA) or, in the absence of SLA data, Local Government Areas).

Because we used incidence density sampling, the OR estimates the mortality rate ratio. Two-sided P values are reported.

To assess whether the ORs varied by time since the program was implemented, we tested for interaction between screening participation and the year of the case’s death (1995–1997, 1998–2001, 2002–2006). In the early years of the program (1995–1997), many participants were attending BreastScreen for the first time (32). It is possible
that women who died soon after the introduction of screening had reduced opportunity for screening before their cancers became clinically detectable, and therefore, less opportunity to benefit from screening compared with women who died later, reducing the effect of screening on mortality during this period. By 2002, the program had been in existence for 7 years—the time period required before mortality benefits were observed in trials (32). We also tested for interaction between screening participation and 10-year age group to assess whether the ORs varied by age.

**Sensitivity analyses**

Our primary analysis was considered the approach least sensitive to theoretical biases, given the information available. To quantify the impact of potential sources of bias not addressed in our primary analysis or addressed differently by other authors, we executed a series of alternative analytic models in which we varied our treatment of sources of potential selection bias, information bias, and confounding.

**Selection bias.** For the primary analysis, we excluded women screened before the age of 50 years to limit the exposure of interest to screening provided within the BreastScreen target age range and to reduce confounding. Women in our sample who participated in screening before 50 years of age are more likely to be disease-free, because they had had a previous negative screen and, as prior screening behavior predicts future behavior, they are more likely to participate in screening. However, this approach may have introduced selection bias by excluding women at higher long-term risk of breast cancer (because women who chose to participate in the screening program before receiving a direct invitation may have had a higher underlying risk of disease) and excluding women at lower long-term risk of breast cancer (the “worried well”). To quantify this bias, in sensitivity analysis A, we allowed women to have been screened from the age of 40 years.

For our primary analysis, we excluded women diagnosed before the age of 50 years and this may have introduced another form of selection bias that can occur if the sample is restricted according to a minimum age at diagnosis (33) because screened cases are typically diagnosed at a younger age than unscreened cases and are therefore more likely to be excluded. In sensitivity analysis B, we allowed women to have been diagnosed from the age of 40 years (and therefore we allowed women to have been screened from age 40 so that they were eligible for exposure until the reference date).

**Information bias.** In our primary analysis, the reference date was set by the first diagnosis in the case–control set. Others have allowed only the case’s diagnosis date to set the reference date (5, 34–36). This potentially introduces information bias because if a control has a breast cancer diagnosis before the date of their matched case, cases in the matched set can still accumulate screens while the control cannot, resulting in potential underestimation of screening benefit (26, 27). To quantify the impact of this potential bias, in sensitivity analysis C, we allowed only cases to set the reference dates.

**Confounding.** Our primary analysis might have been biased because it did not have information on potential confounders other than socioeconomic status and remoteness from health services. As a sensitivity analysis, we used the method of Greenland and Lash (37) to externally adjust for 2 potential sources of confounding: use of hormone replacement therapy and family history of breast cancer (sensitivity analysis D and E). For each confounder, this method requires the estimated prevalence for screened and unscreened women and the breast cancer mortality rate ratio. We used published population estimates of the prevalence of family history of breast cancer and hormone therapy use in the source population and in screened women and observed screening participation rates to derive estimates for prevalence in unscreened women. We estimated maximal plausible values for the mortality rate ratios from the literature. Specifically, for hormone therapy use, from observed data we assumed that 34.6% of BreastScreen participants aged 50 to 69 years used hormone therapy at the time of screening (38) and that 32.9% of all Australian women were current users of hormone therapy (39). We assumed screening participation in this age group was 53% (based on 2000–2001 data for Western Australia; ref. 38). We then inferred that hormone therapy use in unscreened women was 31.0%. The estimated breast cancer mortality rate ratio for hormone therapy users compared with nonusers was 1.48 (40). For family history of breast cancer, we assumed 16.1% of attendees aged 50–69 years had a strong family history of breast cancer (a first degree relative with breast cancer; ref. 38), and 12.0% of all women had a family history (41). Assuming screening participation of 53% as above, we then inferred that the prevalence of family history of breast cancer in unscreened women was 7.4%. The estimated breast cancer mortality rate ratio for women with a family history of breast cancer compared with other women was 2.32 (42).

**Replication of a previous Australian case–control study.** Finally, for comparison with the only other Australian case–control study, we replicated as closely as possible the primary analysis of Roder and colleagues (5) by defining the reference date to be the date of diagnosis of the case, including women screened before the age of 50 years, including women with symptoms at screening and excluding matched sets where cases died before 2002 (sensitivity analysis F).

**Meta-analysis**

We conducted a meta-analysis to test the hypothesis that participation in mammographic service screening programs is associated with a reduction in breast cancer mortality. We included studies that met these criteria: case–control study of mammographic screening with breast cancer death as outcome; evaluation of population mammographic screening programs (not trials); exposure defined as ever versus never screened; exposure...
measured over at least 4 years up to the reference date (to allow reasonable opportunity for exposure to screening for cases and controls); controls matched to cases on age or analysis adjusted for age; and study population age range encompassing or overlapping substantially with our age group of interest (50 to 69 years).

Two investigators (C. Nickson and K. Mason) searched the MEDLINE database through the Thomson Reuters (ISI) Web of Knowledge portal on May 30, 2012, using the search term “breast cancer AND screening AND (mortality OR death) AND (case–control OR case-referent),” with no restrictions on the year published. Studies published in languages other than English were excluded. We searched reference lists of included studies and review articles identified during the search. Search strategies and interim findings are detailed in the Supplementary Data.

Nine studies remained for inclusion in the meta-analysis (refs. 5, 34–36, 43–47; Table 1). Where studies reported multiple estimates derived from different approaches to minimizing bias, C. Nickson and K. Mason jointly identified the analysis most similar to the primary analysis in our own case–control study—an approach we considered least prone to bias. We used ORs not adjusted for “self-selection bias” (confounding due to higher participation by women at increased risk of breast cancer; refs. 34–36, 45), as the need for such adjustment is uncertain (5, 35) and adjustment methods varied. We pooled study ORs using a random effects meta-analysis in Stata (“metan”); heterogeneity was quantified using I^2 (48).

To assess whether country differences biased the pooled estimate, we conducted a meta-regression of ORs by country. We examined publication bias using Egger’s regression asymmetry test for publication bias (49). As a sensitivity analysis, we determined the minimum OR required in an additional study to change the pooled OR CI to include the null effect; we assumed the SE for the additional study’s OR was half that of the smallest SE in the 10 studies included in the meta-analysis.

**Ethics.** Ethical approval for this study was granted by the University of Melbourne Health Sciences Ethics Sub-Committee (0830631) and the Department of Health Western Australia Human Research Ethics Committee (#2009/06).

**Results**

**Primary analysis**

We identified 501 breast cancer deaths (cases) from 1995 to 2006. We excluded 66 of these cases because they had participated in screening before the age of 50 years and a further 7 cases with no data to determine relative disadvantage or remoteness, leaving 428. Each of these had 10 controls, but we excluded an additional 958 controls because they met one or both of the above exclusion criteria or their matched case was excluded. After excluding the controls, one case had no remaining controls and was thus also excluded.

The remaining 427 case–control sets in the primary analysis included an average of 8.5 controls per case (range 2–10) and for 382 (89%) sets, the reference date was the case’s date of diagnosis. Key study variables for cases and controls are summarized in Table 2.

Screening was more common for controls (2,051 of 3,650; 56%) than cases (167 of 427; 39%). The adjusted OR from the primary analysis was 0.48 (95% CI, 0.38–0.59; P < 0.001; Table 3). The ORs varied little by reference age group or year of death and neither interaction was significant (Table 3).

**Sensitivity analyses**

The sample size varied from 123 to 248 case–control sets between sensitivity analyses according to the various inclusion and exclusion criteria. No sensitivity analyses made any material difference to the ORs, which were all between 0.45 and 0.53 (Table 4).

**Meta-analysis**

The ORs from the 10 studies ranged from 0.35 to 0.65 (Fig. 1). The meta-analysis including our own study gave a pooled OR of 0.51 (95% CI, 0.46–0.55) with no significant heterogeneity between studies (I^2 = 0.0%, \( \chi^2 = 8.62, P = 0.473 \)). Excluding our study from the meta-analysis yielded an OR of 0.51 (95% CI, 0.46–0.56).

Meta-regression showed that the pooled OR did not vary significantly by country. There was no evidence of publication bias due to study size (Egger test \( P = 0.970 \)). The minimum OR required in an additional study to produce a pooled estimate with a CI including the null would be 2.4 (95% CI, 2.2–2.6), which would result in a pooled OR = 0.59 (95% CI, 0.34–1.01).

**Discussion**

Participation in the Western Australia BreastScreen program in the target age range of 50 to 69 years was associated with an estimated 52% lower breast cancer mortality. The meta-analysis of results from this and 9 other studies in various settings indicates that, on average, screening participation was associated with a 49% lower mortality from breast cancer.

In practice, potential sources of bias appeared to have had little influence on the results as the ORs were similar in all the sensitivity analyses. Our results were strikingly similar to those of the previous case–control evaluation of the BreastScreen Australia program (OR, 0.52; 95% CI, 0.39–0.69). We found no evidence that the OR varied by year of death, despite recommendations that case–control studies should exclude early years of screening and/or be confined to long-term outcomes since the introduction of screening (32, 35).

The strengths of our study include rigorous linkage of good quality population data from a well-organized population screening program and cancer registry, careful minimization of bias in our primary analysis, and a range.
Table 1. Overview of studies included in the meta-analysis

<table>
<thead>
<tr>
<th>Study</th>
<th>Screening program location</th>
<th>Screening program details</th>
<th>Case selection</th>
<th>Controls</th>
<th>Exposure period in each case–control set</th>
<th>Period for diagnoses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allgood 2008 (56)</td>
<td>United Kingdom (East Anglia)</td>
<td>3-yearly invitations to screening for women aged 50–70; est. 1989</td>
<td>Aged 50–70 years at diagnosis; died 1995–2004</td>
<td>Age-matched and alive when case died</td>
<td>Screened before case diagnosis</td>
<td>1995–2004</td>
</tr>
<tr>
<td>Palli 1989 (60)</td>
<td>Italy (Florence district)</td>
<td>Approx. 30-month intervals, screening invitations to women aged 40–70; est. 1970</td>
<td>Aged ≥ 40 at diagnosis; died 1977–1987</td>
<td>Age- and area-matched and alive when case died</td>
<td>Screened before case diagnosis</td>
<td>1977–1987</td>
</tr>
<tr>
<td>Puliti 2008 (61)</td>
<td>Italy (various regions)</td>
<td>2-yearly invitations to screening for women aged 50–69; est. 1990–1999 (multiple regions; most by 1995)</td>
<td>Aged 50–74 years between program commencement and diagnosis; died 1990–2002</td>
<td>Age-matched and alive when case died; cancer-free when case diagnosed</td>
<td>Case screened before diagnosis; controls screened up to one year after case diagnosis</td>
<td>1990–2001</td>
</tr>
<tr>
<td>Roder 2006 (62)</td>
<td>Australia (South Australia)</td>
<td>2-yearly invitations to screening for women aged 50–69 (but available for women 40–69 and 70+); est. 1994</td>
<td>Aged 45–80 years at death; diagnosed from 1994; died 2002–2000</td>
<td>Age- and area-matched and alive when case died</td>
<td>Screened before case diagnosis</td>
<td>1994–2005</td>
</tr>
<tr>
<td>van Schoor 2011 (63)</td>
<td>Netherlands (Nijmegen)</td>
<td>2-yearly invitations to screening, initially for women aged 35+, then later 50–69; est. 1974</td>
<td>Aged 50–69 years at screening invitation; died 1975–2008</td>
<td>Age- and area-matched and alive when case died; aged 50–69 years at screening invitation</td>
<td>Screened in the 4-year period before the invitation that preceded case diagnosis</td>
<td>1975–2008</td>
</tr>
<tr>
<td>Otto 2012 (64)</td>
<td>Netherlands (Southwest region)</td>
<td>2-yearly invitations to screening for women aged 50–75 (50–69 only in early years); initiated 1990, fully underway 1998</td>
<td>Aged 50–77 at diagnosis; died 1995–2003</td>
<td>Age-matched and alive when case died; cancer-free when case diagnosed</td>
<td>Screened up to a maximum of two invitations prior to the screen preceding case diagnosis</td>
<td>1990–2003</td>
</tr>
</tbody>
</table>
of sensitivity analyses which identified that, in practice, previously described potential biases in the case–control design might have negligible influence on results.

It was not possible for us to address all potential biases. We had limited information on confounders, although external adjustment for 2 possible confounding variables (family history and hormone therapy use) did not change the OR, as found by others (5, 50). This suggests that residual confounding is unlikely to arise from other, unmeasured determinants of screening and mortality (including “self-selection bias”; ref. 1), in keeping with evidence that the baseline risk of death does not differ according to participation (5, 50, 51).

Although we had no information on screening outside the BreastScreen program, rates are estimated to be low (1%–4%; ref. 21).

For the meta-analysis, our selection criteria ensured that each included study had comparable opportunity for exposure between cases and controls, and suitable matching for, or adjustment by, age. The pooled estimate is made more convincing by the lack of heterogeneity and confirmation that study size was not associated with the OR.

The meta-analysis may be subject to publication bias. Only 5 countries are represented, although at least 19 countries have population mammographic screening programs (52). Our shortlist included studies from the United States, but these were excluded because they did not evaluate whole population programs. Our restriction to papers published in English led to the exclusion of only 2 papers, (53, 54) both from the Netherlands, which is already well represented in the meta-analysis. While studies may have been conducted but not published due to a null finding, this is unlikely as this is a contested field and ecological studies that found no association have been published and received considerable attention (8, 11). Finally, it is possible that some studies were missed due to the search terms used. However, our assessment of 3 previous review papers revealed no additional studies for inclusion, and our sensitivity analysis showed only a study with an

### Table 2. Summary figures for cases and controls included in the analysis

<table>
<thead>
<tr>
<th></th>
<th>Cases</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>N in sample for analysis</td>
<td>427</td>
<td>3,650</td>
</tr>
<tr>
<td>Age information (y): mean (SD); range</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age at first screen (if ever screened)</td>
<td>56.5 (4.7); 50–70</td>
<td>57.7 (5.2); 50–79</td>
</tr>
<tr>
<td>Age at reference date (first diagnosis in set)</td>
<td>59.6 (5.6); 50–69</td>
<td>60.4 (5.4); 50–69</td>
</tr>
<tr>
<td>Age at death</td>
<td>63.5 (6.0); 50–77</td>
<td>—</td>
</tr>
<tr>
<td>Calendar year information: mean (range)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relative socioeconomic disadvantage (SEIFA score): mean (SD)</td>
<td>1,004.5 (89.9)</td>
<td>1,009.1 (90.0)</td>
</tr>
<tr>
<td>Remoteness from services (ARIA): distribution</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Major city</td>
<td>73.3%</td>
<td>74.1%</td>
</tr>
<tr>
<td>Inner regional</td>
<td>11.2%</td>
<td>13.0%</td>
</tr>
<tr>
<td>Outer regional</td>
<td>10.1%</td>
<td>8.2%</td>
</tr>
<tr>
<td>Remote</td>
<td>3.0%</td>
<td>3.3%</td>
</tr>
<tr>
<td>Very remote</td>
<td>2.3%</td>
<td>1.4%</td>
</tr>
</tbody>
</table>

### Table 3. ORs for death from breast cancer according to history of participation in the BreastScreen Western Australia program, overall and according to year of death and age at reference date

<table>
<thead>
<tr>
<th>Analysis</th>
<th>OR (95%CI)*</th>
<th>P</th>
<th>P for interaction term</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aged 50–69, cases died 1995–2006</td>
<td>0.48 (0.38–0.59)</td>
<td>&lt;0.001</td>
<td>—</td>
</tr>
<tr>
<td>Year of death</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cases died 2002–2006</td>
<td>0.49 (0.37–0.66)</td>
<td>&lt;0.001</td>
<td>0.841</td>
</tr>
<tr>
<td>Cased died 1998–2001</td>
<td>0.43 (0.29–0.64)</td>
<td>&lt;0.001</td>
<td>—</td>
</tr>
<tr>
<td>Cases died 1995–1997</td>
<td>0.51 (0.25–1.07)</td>
<td>0.077</td>
<td>—</td>
</tr>
<tr>
<td>Age at reference date</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age 50–59 at reference date</td>
<td>0.52 (0.37–0.72)</td>
<td>&lt;0.001</td>
<td>0.507</td>
</tr>
<tr>
<td>Age 60–69 at reference date</td>
<td>0.44 (0.33–0.59)</td>
<td>&lt;0.001</td>
<td>—</td>
</tr>
</tbody>
</table>

*Adjusted for remoteness and relative socioeconomic disadvantage
extreme finding (OR = 2.4) could influence the pooled estimate CI to include the null effect.

Meta-regression according to country showed no evidence of any association between country and estimated mortality reductions associated with screening, although this result should be interpreted with caution due to low statistical power. If the finding does hold, it would be reasonable to generalize to other countries with screening programs similar to those in our evaluation (Table 1), with consideration of underlying population health profiles and the availability of treatment following diagnosis.

Unlike case–control studies of screening, inconsistent results have been reported from ecological studies. The inconsistency in results of ecological studies is partly due to differences in measuring screening activity (17).

Table 4. ORs for death from breast cancer according to history of participation in the BreastScreen Western Australia program, for primary analysis and a range of sensitivity analyses

<table>
<thead>
<tr>
<th>Analysis</th>
<th>n</th>
<th>Cases (deaths)</th>
<th>Controls</th>
<th>Mean (SD) age at reference date</th>
<th>OR (95% CI)*</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary analysis</td>
<td>4,077</td>
<td>167/427 (39.1)</td>
<td>2,051/3,850 (56.2)</td>
<td>60.3 (5.5)</td>
<td>0.48 (0.38–0.59)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Sensitivity analysis A</td>
<td>5,273</td>
<td>230/490 (46.9)</td>
<td>2,957/4,783 (61.8)</td>
<td>58.8 (5.8)</td>
<td>0.51 (0.42–0.62)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Sensitivity analysis B</td>
<td>6,271</td>
<td>248/582 (42.6)</td>
<td>3,176/5,689 (55.8)</td>
<td>56.9 (6.8)</td>
<td>0.53 (0.44–0.63)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Sensitivity analysis C</td>
<td>4,024</td>
<td>165/424 (38.9)</td>
<td>2,064/3,800 (57.3)</td>
<td>60.4 (5.5)</td>
<td>0.45 (0.36–0.56)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Sensitivity analysis D</td>
<td>4,077</td>
<td>0.49</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity analysis E</td>
<td>4,077</td>
<td>0.45</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sensitivity analysis F</td>
<td>2,499</td>
<td>123/232 (53.0)</td>
<td>1,532/2,267 (67.6)</td>
<td>58.9 (5.5)</td>
<td>0.52 (0.39–0.69)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

NOTE: The primary analysis includes case–control sets where the case died between 1995 and 2006 and the reference date (first breast cancer diagnosis in the case–control set) occurred at age 50–69 years and between 1995 and 2006. Women were defined as exposed to screening participation if they participated at least once between the age of 50 years and the reference date. Sensitivity analyses vary in sample size according to inclusions or exclusions for each analysis (see Methods).

*Adjusted for remoteness and relative socioeconomic disadvantage

bMethod does not allow for meaningful values in other columns, or CIs for ORs.
Geographic approaches include comparing rates of participation and mortality within and between municipalities (6), counties (16), or countries (8). The smaller and more homogenous the geographic area, the more likely it is that the participation rate is applicable to all women in the area and the less likely there is to be confounding due to regional differences in factors such as breast cancer treatment (e.g., adjuvant therapy).

For the two small-area Australian ecological studies (6, 7), the first (restricted to New South Wales) estimated that a 60% participation rate in BreastScreen was associated with a 29% lower breast cancer mortality (relative risk 0.71; 95% CI, 0.60–0.85; ref. 6), while in the second, national, study, a 60% participation rate was associated with a 22% lower mortality (relative risk, 0.78; 95% CI, 0.69–0.88; ref. 7). In the second study in particular, several potential sources of confounding within small geographic areas were included in the analysis. In contrast, the European ecological study that drew substantial media attention compared mortality trends for 3 countries (11) and while the authors claimed the countries were similar with respect to health care and other determinants of breast cancer mortality, the analysis has been criticized by screening experts (55); indeed, interpretation of country-level time trends is prone to confounding by changes in treatment and other secular changes that are likely to vary by country.

Examining time trends by age within populations is another common but weak ecological design that is subject to confounding. For example, in the aforementioned Australian study the authors concluded that BreastScreen was ineffective because the decline in breast cancer mortality from 1991 to 2007 was greatest for women aged 40 to 49 at death rather than the target age groups for screening (8). The observed decline could be explained by other factors such as age differences in treatment access and efficacy, so it is inappropriate to infer that screening is ineffective for other age groups.

Our findings suggest that screening is of benefit to women who choose to be screened. We estimate a 52% reduction in breast cancer mortality for women who participate in screening in Australia, and an average 49% reduction from the studies included in our meta-analysis. Sensitivity analyses to address potential biases of case–control studies showed that none of the assumptions had any material effect, thus showing that these potential biases might be overstated. Our results were consistent with those of other case–control studies that have taken a range of approaches to minimizing bias. While ecological studies of large geographical areas may be sensational—particularly when they find no effect—interpreting trends in mortality is complex. Meanwhile, case–control studies offer a robust and consistent contribution to growing evidence that screening is of benefit to women who choose to be screened.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors’ Contributions

Conception and design: C. Nickson, D.R. English, A.M. Kavanagh
Development of methodology: C. Nickson, K.E. Mason, D.R. English, A.M. Kavanagh
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): C. Nickson
Analysis and interpretation of data (e.g., statistical analysis, bios-statistics, computational analysis): C. Nickson, K.E. Mason, D.R. English
Writing, review, and/or revision of the manuscript: C. Nickson, K.E. Mason, D.R. English, A.M. Kavanagh
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): C. Nickson, K.E. Mason
Study supervision: A.M. Kavanagh

Acknowledgments

The authors thank the Data Linkage Branch of the Department of Health WA for enabling high-quality linkage of health datasets. The authors also thank Associate Professor Elizabeth Wylie, Medical Director of BreastScreen WA for supporting the inaugural linkage of the BreastScreen WA dataset by the Data Linkage Branch for the purpose of this study.

Grant Support

The work was supported by a Project Grant from the National Health and Medical Research Council of Australia (grant number 509357).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received April 18, 2012; revised June 19, 2012; accepted July 5, 2012; published online September 6, 2012.

References


# GENERAL ULTRASOUND

What are some *INDICATIONS* for each study?

<table>
<thead>
<tr>
<th>Study</th>
<th>Indications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abdominal US</td>
<td></td>
</tr>
<tr>
<td>Pelvic US</td>
<td></td>
</tr>
<tr>
<td>Thyroid US</td>
<td></td>
</tr>
<tr>
<td>Carotid US</td>
<td></td>
</tr>
</tbody>
</table>

What are some *LIMITATIONS* of each study

<table>
<thead>
<tr>
<th>Study</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abdominal US</td>
<td></td>
</tr>
<tr>
<td>Pelvic US</td>
<td></td>
</tr>
<tr>
<td>Thyroid US</td>
<td></td>
</tr>
<tr>
<td>Carotid US</td>
<td></td>
</tr>
</tbody>
</table>
Are there any specific instructions you should discuss with the patient prior to the exam?

<table>
<thead>
<tr>
<th>Modality</th>
<th>Instructions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abdominal US</td>
<td></td>
</tr>
<tr>
<td>Pelvic US</td>
<td></td>
</tr>
<tr>
<td>Thyroid US</td>
<td></td>
</tr>
<tr>
<td>Carotid US</td>
<td></td>
</tr>
</tbody>
</table>

**Recommended Reading:**

NUCLEAR MEDICINE

Background Information:

1. How are nuclear medicine studies different from x-ray, ultrasound, CT and MRI studies?

2. How are nuclear medicine studies used in patient care? Diagnosis? Treatment? Both?

3. How are radiopharmaceutical agents introduced into the body?

4. Are radiopharmaceutical agents safe?

What are some INDICATIONS for each study?

Bone Scan

BMD

Renal Scan
Thyroid Scan

Exercise Cardiac Stress Test

Pharmacological Cardiac Stress Test

Cases:

1. 51 yo M patient with possible chronic cholecystis. What test would you order?

2. 40 yo F with persistent right upper quadrant pain after a cholecystectomy 1 year ago. What test would you order?

3. 60 yo M with a history of prostate cancer. New onset of R hip pain. What test(s) would you order?

4. 28 yo F, marathon runner. New onset of pain in the right lower leg. Initial x-rays are negative. What test would you order if you are concerned about a stress fracture?
5. 25 yo F who is breastfeeding and requires a thyroid scan. She is asking about the effects of the radiopharmaceutical on her breast milk? What would you tell her?

Recommended Reading:
http://interactive.snm.org/docs/whatisnucmed2.pdf (from the Society of Nuclear Medicine)
**POST-ACE SELF-EVALUATION**

*(General DI)*

1. **How would you rate your overall knowledge of General DI in the community after the ACE?**

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No knowledge</td>
<td>Excellent knowledge</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2. **Did you achieve the 3 goals you identified before the General DI ACE?**

   ________________________________

3. **What are 3 new things you learned from the General DI ACE?**

   a) ________________________________

   b) ________________________________

   c) ________________________________

4. **How will you change your practice after completing the General DI ACE?**

   ________________________________

   ________________________________

   ________________________________
CanMEDS EVALUATION  
*(General DI)*

Circle the non-Medical Expert CanMEDS Roles* addressed during the General DI ACE and give an example of how each Role was addressed.

*Adapted from 2005 The Royal College of Physicians and Surgeons of Canada* 
(http://www.royalcollege.ca/shared/documents/canmeds/the_7_canmeds_roles_e.pdf)
OVERALL EVALUATION OF ACE
(General DI)

1. What is your overall impression of the General DI ACE?

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very Poor</td>
<td>Poor</td>
<td>Average</td>
<td>Very Good</td>
<td>Excellent</td>
</tr>
</tbody>
</table>

2. What were the 3 most helpful elements of the General DI ACE?

a) ___________________________________________________________

b) ___________________________________________________________

c) ___________________________________________________________

3. What are 3 elements of the General DI ACE that were the least helpful?

a) ___________________________________________________________

b) ___________________________________________________________

c) ___________________________________________________________

4. Is there anything that you would like to be included in future General DI ACE’s?

_________________________________________________________________

_________________________________________________________________

_________________________________________________________________

5. Additional Comments?

_________________________________________________________________

_________________________________________________________________
Image-Guided Pain Management/MSK ACE

(Advanced Spinal Care Clinic—ASCC)
**PRE-ACE SELF-EVALUATION**
*(Image-guided Pain Management/MSK at ASCC)*

1. How would you rate your overall knowledge of Image-guided Pain Management & Musculoskeletal Imaging in the community? (i.e. Types of procedures? Types of studies? Indications? Limitations?)

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No knowledge</td>
<td>Excellent knowledge</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2. What are 3 goals you would like to achieve during the Image-guided Pain Management/MSK ACE?

a) ____________________________

b) ____________________________

c) ____________________________

3. What are 3 specific questions you would like to have answered during the Image-guided Pain Management/MSK ACE?

a) ____________________________

b) ____________________________

c) ____________________________
IMAGE-GUIDED PAIN MANAGEMENT & MUSCULOSKELTAL IMAGING

Background Information:

1. What types of musculoskeletal imaging is available in community radiology clinics?

2. What does “Image-guided Pain Management” mean with respect to radiology interventions?

3. Are musculoskeletal procedures performed in community radiology clinics covered by Alberta Health?

4. What are some indications for imaging in acute low back pain?

5. What are some indications for imaging in traumatic knee pain?

6. What are some indications for imaging in non-traumatic knee pain?

Recommended Reading:
IMAGE-GUIDED PAIN MANAGEMENT & MUSCULOSKELTAL IMAGING

Background Information:

1. What types of musculoskeletal imaging is available in community radiology clinics?

2. What does “Image-guided Pain Management” mean with respect to radiology interventions?

3. Are musculoskeletal procedures performed in community radiology clinics covered by Alberta Health?

4. What are some indications for imaging in acute low back pain?

5. What are some indications for imaging in traumatic knee pain?

6. What are some indications for imaging in non-traumatic knee pain?

Recommended Reading:
POST-ACE SELF-EVALUATION  
(Image-guided Pain Management/MSK at ASCC)

1. How would you rate your overall knowledge of Image-guided Pain Management & MSK Imaging in the community after the ACE?

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No knowledge</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Excellent knowledge</td>
</tr>
</tbody>
</table>

2. Did you achieve the 3 goals you identified before the Image-guided Pain Management/MSK Imaging ACE?

________________________________________________________________________

3. What are 3 new things you learned from the Image-guided Pain Management/MSK Imaging ACE?

a) ______________________________________________________________________

b) ______________________________________________________________________

c) ______________________________________________________________________

4. How will you change your practice after completing the Image-guided Pain Management/MSK Imaging ACE?

________________________________________________________________________

________________________________________________________________________
CanMEDS EVALUATION
(Image-guided Pain Management/MSK at ASCC)

Circle the non-Medical Expert CanMEDS Roles* addressed during the Image-guided Pain Management & MSK Imaging ACE and give an example of how each Role was addressed.

CanMEDS 2005 FRAMEWORK

*Adapted from 2005 The Royal College of Physicians and Surgeons of Canada
(http://www.royalcollege.ca/shared/documents/canmeds/the_7_canmeds_roles_e.pdf)
OVERALL EVALUATION OF ACE
(Image-guided Pain Management/MSK at ASCC)

1. What is your overall impression of the Image-guided Pain Management & MSK Imaging ACE?

   1 2 3 4 5
   Very Poor Poor Average Very Good Excellent

2. What were the 3 most helpful elements of the Image-guided Pain Management & MSK Imaging ACE?
   a) 
   b) 
   c) 

3. What are 3 elements of the Image-guided Pain Management & MSK Imaging ACE that were the least helpful?
   a) 
   b) 
   c) 

4. Is there anything that you would like to be included in future Image-guided Pain Management & MSK Imaging ACE’s?

   ____________________________________________________________
   ____________________________________________________________
   ____________________________________________________________

5. Additional Comments?

   ____________________________________________________________
   ____________________________________________________________
Low back pain is one of the most common reasons for visits to physicians in the ambulatory care setting. Estimated medical expenditures related to low back pain have increased disproportionately relative to the more modest increase in the prevalence of self-reported low back pain in the past decade. The increase in spine care expenditures has not been associated with improved patient outcomes. Evidence-based order templates presented in this article are designed to assist practitioners through the process of managing patients with acute low back pain. A logical method of choosing, developing, and implementing clinical decision support interventions is presented that is based on the best available scientific evidence. These templates may be reasonably expected to improve patient care, decrease inappropriate imaging utilization, reduce the inappropriate use of steroids and narcotics, and potentially decrease the number of inappropriate invasive procedures.

**Key Words:** Acute low back pain, computerized decision support, clinical practice guidelines, imaging utilization, computerized order entry systems

**BACKGROUND**

Low back pain is one of the most common reasons for visits to physicians in the ambulatory care setting [1]. In one study, 26.4% of adults reported episodes of acute low back pain within the past 3 months [2]. Although the prevalence of self-reported low back pain has increased only modestly in the past decade, estimated medical expenditures related to back pain have increased substantially [3]. Increased utilization of medical imaging is one component of this cost increase [4,5]. The total costs related to back pain, both direct and indirect, are estimated to be >$100 billion per year in the United States [6]. Despite the increase in overall spine care expenditures associated with medical imaging, there has not been an incremental improvement in patient outcomes [3,7].

The approach to the workup and management of low back pain by physicians and other practitioners is inconsistent. There is significant variability in the diagnostic workup of back pain among physicians within and between specialties [8,9]. Survey data indicate that there is little consensus among physicians regarding what treatments are effective for low back pain [10]. Despite mixed evidence for the efficacy of surgical intervention in different types of low back pain, rates of back surgery in the United States have been estimated to be >40% higher than in other developed countries [11]. Furthermore, there seems to be a relationship between the increasing use of advanced imaging and accelerating rates of surgical intervention [12].

**EVIDENCE-BASED CLINICAL PRACTICE GUIDELINES**

Recent efforts have been made to synthesize and summarize the extensive and sometimes confusing literature on the evaluation and management of low back pain [7,13-15]. Clinical practice guidelines have been published in the United States [16,17] and abroad [18], with the aim of decreasing variability, improving the quality of care, increasing patient safety, and encouraging medical care that is based on the best available scientific evidence. Despite differences in culture, local regional trends, and health care systems, there is remarkable similarity in the various low back pain clinical practice guidelines [18]. Although there is seemingly a broad international consensus among the authors of clinical practice guidelines, significant gaps exist between current clinical practices and evidence-based recommendations.
There are well-known barriers to the widespread implementation of evidence-based clinical practice guidelines [19-21]. Common barriers include lack of awareness of the guidelines, belief that the guidelines will not produce the desired results, disagreement with the guidelines, belief that guidelines cannot be effectively implemented, and the inability to overcome the inertia of previous practice. In the case of low back pain, patient expectations, miscommunication, and factual inaccuracies held by physicians have been reported as reasons for deviation from clinical practice guidelines [22,23]. Even when practice guidelines are well understood and generally followed, there are certain situations in which practitioners deviate from the recommendations. For instance, the presence of sciatica has been associated with divergence from low back pain practice guidelines [24]. Most of the efforts to better align everyday practice with evidence-based low back pain guidelines have focused on educational outreach [25,26]. More recent efforts in this regard are taking place at the point of care with the use of technology.

**CLINICAL DECISION SUPPORT AND MEANINGFUL USE**

Clinical decision support (CDS) systems, in general terms, are software applications designed to assist health care providers in decision making throughout the health care process. When used at the order entry stage, these applications provide a unique opportunity to marry evidence-based clinical guidelines with computerized physician order entry systems. Clinical decision support interventions have been in existence for decades, yet there is a lack of widespread adoption in the United States. Recently, the federal government has provided monetary incentives for the implementation of CDS systems.

The Health Information Technology for Economic and Clinical Health Act of the American Recovery and Reinvestment Act of 2009 outlines a set of incentive payments for physicians and hospitals that demonstrate meaningful use of health IT. At the outset, the meaningful use incentive rule requires the implementation of one CDS system “relevant to specialty or high clinical priority, including for diagnostic test ordering, along with the ability to track compliance” [27]. In the wake of this legislation, providers have been slowly mobilizing to meet meaningful use standards.

In this article, we present a framework for the development of acute low back pain decision support tools with standardized order sets. This framework can be used by medical providers and systems designers to develop decision support applications that are customized to their unique practice setting. Order sets created using evidence-based best practices should improve clinical performance, establish a standard of care for an institution, streamline patient encounters, encourage regulatory compliance, and achieve the ultimate goals of better quality of care and outcomes.

Although the focus of this article is on developing CDS tools for use within the computerized physician order entry environment, this framework can also be used by providers that do not intend to meet new federal standards but have a desire to incorporate evidence-based low back pain order sets and evidence-based standards for imaging utilization into their clinical practice.

**CLINICAL DECISION SUPPORT TEMPLATES FOR ACUTE LOW BACK PAIN**

Patients with acute low back pain (symptoms lasting <4 weeks) first undergo a thorough history and physical examination, after which they are placed into 1 of 3 broad clinical categories: nonspecific low back pain, low back pain potentially associated with radiculopathy or spinal stenosis, or low back pain potentially associated with a specific cause (Table 1). Order set templates have been devised for the initial visit and follow-up visits for patients falling into each of these clinical categories.

The reader will notice that the order sets do not contain specifics with regard to medication dose, route, or frequency. This is an intentional omission and is due to the number of variables that come into play when prescribing these medications, such as renal function, hepatic function, allergies, and so on. These data could be input manually, with the use of drop-down lists, or integrated into other medication-related decision support programs to aid clinicians in choosing the safest, most effective medication for each clinical scenario. The reader will also notice that each order set includes an input field titled “other.” These fields allow for template customization to incorporate local practice preferences.

The order sets provided herein are presented as one possible type of decision support tool among several possible tools that could be developed and used in this clinical setting. For instance, information buttons

<table>
<thead>
<tr>
<th>Table 1. Adult with low back pain (acute)</th>
</tr>
</thead>
<tbody>
<tr>
<td>- History and physical key points</td>
</tr>
<tr>
<td>- Duration and nature of symptoms</td>
</tr>
<tr>
<td>- Presence of red flags (trauma history, unintentional weight loss, immunosuppression, history of cancer, intravenous drug use, steroid use, osteoporosis, age &gt; 50 y, focal neurologic deficit, progression of symptoms)</td>
</tr>
<tr>
<td>- Symptoms of spinal stenosis, radiculopathy</td>
</tr>
<tr>
<td>Decision point (&lt;4 weeks of symptoms)</td>
</tr>
<tr>
<td>- No red flags, signs, or symptoms of spinal stenosis or radiculopathy</td>
</tr>
<tr>
<td>- Go to order set for nonspecific acute low back pain</td>
</tr>
<tr>
<td>- Signs or symptoms of spinal stenosis or radiculopathy</td>
</tr>
<tr>
<td>- Go to order set for acute low back pain with radiculopathy or spinal stenosis</td>
</tr>
<tr>
<td>- Red flags present</td>
</tr>
<tr>
<td>- Go to order set for acute low back pain with red flags</td>
</tr>
</tbody>
</table>
could be used to provide real-time access to current clinical practice guidelines. Alerts and reminders during the order process may suggest alternative diagnostic or treatment modalities on the basis of patientspecific data. A more structured CDS system may be incorporated in which the end user is required to affirm the presence of certain criteria (ie, "red flags") before ordering an MRI examination or provide justification before prescribing steroids or narcotics. The order sets can be programmed to initiate the printout of patient handouts including information on low back pain, descriptions of self-care techniques, and the date and time of the follow-up visit.

Evidence-based practice guidelines for the evaluation and management of chronic low back pain (>3 months) overlap somewhat with acute low back pain guidelines. However, important differences exist that preclude adequate coverage of chronic low back pain evaluation and management in this paper.

**Nonspecific Low Back Pain**

Nonspecific low back pain is predominantly localized to the back and is not associated with signs or symptoms of the entities found in the other two categories (ie, red flags, specific conditions, radiculopathy, or spinal stenosis). Nonspecific low back pain is not associated with significant functional impairment or rapidly progressive neurologic deficits.

Self-care techniques with established efficacy in this group include advising patients to remain active, the use of patient handouts, and the application of superficial heat [16]. There is good evidence that nonsteroidal anti-inflammatory drugs and skeletal muscle relaxants are effective for short-term pain relief. There is fair evidence that acetaminophen, tramadol, and benzodiazepines are effective for short-term pain relief. Although less effective, acetaminophen should be considered as a first-line medication because of its superior side effect profile. Opioids and tramadol may be prescribed for patients with severe, debilitating pain that is not responsive to acetaminophen or nonsteroidal anti-inflammatory drugs [16]. Referrals for physical or occupational therapy may also be considered. Imaging and invasive interventions are not recommended at this stage [7,14,16]. The evidence-based order form for this clinical pathway is found in Figure 1.

The patient is reassessed in 4 weeks. If there is improvement in the patient’s symptoms at follow-up, educational materials are provided, and instructions on self-care are reinforced. Referrals for physical therapy, occupational therapy, or cognitive-behavioral therapy may be considered at this point, and further follow-up may be scheduled as is appropriate (Figure 2 online).

If there is no improvement in symptoms and there continues to be no evidence of red flags, radiculopathy, or spinal stenosis, referral for imaging may be considered. MRI is the preferred imaging modality, but CT may also provide useful information in those situations in which MRI is not possible [14,16]. The order set for this clinical category is found in Figure 3 online.

If there are signs and symptoms of radiculopathy or spinal stenosis at the 4-week follow-up visit, imaging is considered only for those patients who are realistic candidates for invasive procedures. At this point, surgical and interventional pain management referrals may be considered. MRI is the modality of choice in this situation, with CT considered a viable option when MRI cannot be performed [14,16]. Figure 3 online contains a complete order set for this clinical pathway.

**Radiculopathy and Spinal Stenosis**

Radiculopathy is defined as nerve root dysfunction manifesting as pain, paresthesia, reduced sensory function, decreased deep tendon reflexes, or weakness. The available evidence is insufficient to make specific self-care recommendations for patients in this group. However, there is evidence suggesting that self-care techniques used for patients with nonspecific low back pain may be safely used in this patient group [16]. Referral for physical or occupational therapy may be considered during the initial visit. Gabapentin has been shown to have small, short-term treatment effects in patients with radiculopathy. Otherwise, there is not sufficient evidence to establish specific medication recommendations for patients in this group. Imaging is not a part of the initial evaluation unless red flags are present [7,14,16], including patients being considered for epidural steroid injections [28]. An evidence-based order set for the initial visit is provided in Figure 4.

Follow-up takes place 4 weeks after the initial visit. If there is clinical improvement, self-care instructions and educational materials are provided, and additional follow-up is scheduled as needed (Figure 5 online). The patient is reassessed for psychosocial factors that may predict poorer long-term outcomes, such as depressive mood, somatization, and distress [29]. If there is no clinical improvement and the patient is a realistic candidate for invasive interventions, a referral for imaging may be appropriate. MRI is the modality of choice, with CT as a second option [14]. Surgical or interventional pain management referrals may be discussed with the patient at this point [16]. An order set for patients with radiculopathy or spinal stenosis who have not improved at the time of the initial follow-up visit is found in Figure 6 online.

**Red Flags**

This category comprises the small percentage of patients who display red flags indicating the possibility of a serious underlying condition, such as malignancy, vertebral infection, vertebral compression fracture, cauda equina syndrome, and ankylosing spondylitis. Also included in this category are patients with severe or progressive neurologic deficits. Indicators of potentially serious underlying causes of acute low back pain...
include recent serious trauma or milder trauma in a patient >50 years of age, immunosuppression, intravenous drug use or abuse, advanced age (>70 years), and osteoporosis [14,16].

The diagnostic workup may include routine laboratory tests (eg, basic metabolic panel, complete blood count), erythrocyte sedimentation rate, C-reactive protein, human leukocyte antigen B27, serum or urine electrophoresis, electromyography, and nerve conduction velocity testing.

MRI is generally considered the initial imaging modality of choice for patients with red flags. CT with or without myelography may be used when there are contraindications to MRI. Plain radiography and \(^{99m}\)Tc bone scans may be considered acceptable modalities for the initial imaging workup of certain patients in this category [14]. Management of patients with red flags consists of treating the underlying etiology. An additional imaging workup may be required for treatment planning when a specific underlying etiology is identified [14]. A complete evidence-based order set for patients presenting with red flags is found in Figure 7.

**DEVELOPMENT AND IMPLEMENTATION OF CLINICAL DECISION SUPPORT SYSTEMS**

Once the decision has been made to develop a CDS system for acute low back pain, stakeholders are assembled to discuss the objectives and desired outcome. Common objectives include improving patient care, reducing patient inconvenience, increasing efficiency, and reducing costs. Once consensus has been reached on the objectives and outcomes, the type of CDS application is determined.

The number of decision support applications has grown significantly over time. These applications may be integrated into the electronic medical record or exist as standalone programs. Most CDS applications are commercial products, but a number of “homegrown” applications have been developed at academic centers. Common decision support applications include reference tools (eg, information buttons, Web searches), order sets, documentation templates,
protocol support, data displays, alerts, and reminders. Integral to the decision of the type of CDS tool is determination of the degree of end user control over whether to launch the tool and whether to follow recommendations generated by the tool. The degree of user control may have a significant impact on the effectiveness of the intervention.

When considering the various CDS tools, it is important to discuss the feasibility of the different options with the systems designer or vendor. Depending on the type of electronic medical record, there may be decision support tools that are already available or CDS components that can be used (eg, interfaces, logic rules, templates). Additional hardware or software may be required to fully implement the desired application. Organizations should also be prepared to allocate additional personnel during CDS implementation and thereafter for ongoing system maintenance.

Kawamoto et al [30] identified 4 features that are critical to the successful deployment of decision support systems: automatic provision of decision support as a part of clinician workflow, provision of recommendations rather than assessments, provision at the time and location of clinical decision making, and computer-based decision support. Those systems that had all 4 features showed higher levels of success. Similar experiences have been reported elsewhere in the literature [31].

Potential barriers to the implementation of CDS systems include real or perceived threats to physician autonomy, harm to the doctor-patient relationship, prior experience with poorly functioning computerized systems, and overreliance on a computer application [32]. If effectively addressed early in the design and implementation process, these factors can be minimized.

---

**Radiculopathy**: Dysfunction of a nerve root associated with pain, sensory impairment, weakness, or diminished deep tendon reflexes in nerve root distribution.

**Spinal Stenosis**: Low back or radicular pain that increases with walking and improves with flexion (sitting or propping)

---

**PHARMACOLOGIC**
- Acetaminophen
- NSAID
- Skeletal muscle relaxants
- Gabapentin
- Other (tramadol, opioids, benzodiazepine)

**ACTIVITY**
- Normal
- Heat application
- Other

**EDUCATION**
- Back Pain Pamphlet (eg, The Back Book*)
- Other

**CONSULTS**
- Physical Therapy
- Occupational Therapy
- Other

**RETURN APPOINTMENT**
- 4 weeks
- Other


**Fig 4.** Acute low back pain with symptoms of spinal stenosis or radiculopathy pathway: initial visit. NSAID = nonsteroidal anti-inflammatory drug.
Decision support systems can be costly to implement and maintain. In addition to the potential capital outlays for new software or hardware, demands on staff time can be substantial. This is the case not only before and during implementation but also on an ongoing basis to keep the decision support tools up to date [33].

Assessment of a CDS intervention properly begins before implementation, with the collection of baseline data. Preimplementation and postimplementation data can then be compared to evaluate program efficacy. Arguments have also been made for small-scale evaluations that take place throughout the design and implementation phases that incorporate feedback from all stakeholders [34].
Focal neurologic deficit with progressive or disabling symptoms, cauda equine symptoms (urinary retention, multilevel motor deficit, fecal incontinence, saddle anesthesia)

**Imaging:**
- □ MRI of the lumbar spine without contrast (preferred)
- □ MRI of the lumbar spine without and with contrast
- □ Myelography and postmyelography CT of the lumbar spine
- □ CT lumbar spine with or without IV contrast
- □ Other

**Other:**
- □ Electromyography/nerve conduction velocity

**PHARMACOLOGIC**
- □ Acetaminophen
- □ NSAID
- □ Antidepressants (TCA†)
- □ Benzodiazepines
- □ Tramadol
- □ Opioids
- □ Other

**ACTIVITY**
- □ Normal
- □ Heat application
- □ Other

**EDUCATION**
- □ Back Pain Pamphlet (eg, The Back Book‡)
- □ Other

**CONSULTS**
- □ Physical Therapy
- □ Occupational Therapy
- □ Psychiatry
- □ Neurology
- □ Neurosurgery
- □ Orthopedics
- □ Hematology-Oncology
- □ Infectious Disease
- □ Endocrine
- □ Other
DISCUSSION

Decision support interventions have the potential to increase clinician speed and efficiency. However, evidence-based order sets have well-known pitfalls that must be taken into consideration [34]. For instance, order sets that are difficult or inconvenient to access may not be used. Once implemented, order sets that are not regularly reviewed and revised quickly become outdated. Additionally, although order sets have many useful features, they typically cannot be sufficiently customized (eg, cannot be adjusted on the basis of current laboratory results, medication lists, or other dynamic factors) [35].

Systematic reviews have demonstrated improvement in clinician performance after CDS interventions [36,37]. A host of CDS systems used in a variety of clinical settings were included in these reviews. Although improvement in physician performance was observed in most of these studies, such was not often the case in terms of patient outcomes. At the present time, it is not clear what CDS features, if any, lead to improved patient outcomes.

A handful of studies have evaluated decision support systems implemented in the diagnostic imaging order entry process. An early study demonstrated that the implementation of decision support resulted in patterns of imaging utilization that more closely resembled imaging expert recommendations. Other groups have reported reductions in the percentage of low-utility imaging examinations [38,39] and reductions in the rate of growth of imaging utilization after the implementation of CDS systems [40].

Blackmore et al [41] developed an evidence-based decision support tool that specifically targeted inappropriate lumbar MRI utilization, the only intervention of its type to date. The CDS intervention in this case consisted of providers answering a series of questions during the order entry process. Providers that did not document compliance with institutionally approved indications were denied access to MRI, although alternatives were provided. This more restrictive intervention resulted in a sustained decrease in lumbar MRI utilization by 23.4%.

In addition to the meaningful use incentives previously discussed, payers in some regions are beginning to incentivize physicians to use CDS systems by waiving preauthorization requirements. In these arrangements, physician orders must fall within a certain range of the established imaging appropriateness criteria to obtain such a waiver. The appeal of such arrangements is that they are more likely to be evidence based, transparent, education based, and to increase the efficiency of payers and physicians alike.

CONCLUSIONS

We have presented a framework for the development of decision support applications for acute low back pain. At the initial visit, patients are categorized into 1 of 3 groups after a thorough history and physical examination: non-specific low back pain, low back pain potentially associated with radiculopathy or spinal stenosis, or low back pain potentially associated with a specific cause. Evidence-based order sets are provided for each category that are intended to guide practitioners through the process of evaluation, management, and follow-up of patients.

Order set templates for use at the initial follow-up visit (4 weeks) provide evidence-based recommendations for appropriate imaging, laboratory workup, referral for invasive procedures, or surgical consultation. Reminders to reassess for psychological factors and red flags can be integrated into the order screen at this stage.

The evidence-based order templates we have presented are designed to assist practitioners with the sometimes confusing process of managing patients with acute low back pain. We have presented a logical method of choosing, developing, and implementing CDS interventions that is based on the best available evidence. A carefully designed CDS system may be reasonably expected to improve patient care, decrease inappropriate imaging utilization, reduce the inappropriate use of steroids and narcotics, and potentially decrease the number of inappropriate invasive procedures. Ideally, these templates could also be used to develop transparent criteria for payer coverage determinations with regard to imaging, medications, procedures, and surgical interventions.
REFERENCES


17. Veterans Health Administration and Department of Defense. VHA/DOD clinical practice guideline for the management of low back pain or sciatica in the primary care setting, version 1.0. Washington, DC: Veterans Health Administration; 1999.


20. Vertens Health Administration and Department of Defense. VHA/DOD clinical practice guideline for the management of low back pain or sciatica in the primary care setting, version 1.0. Washington, DC: Veterans Health Administration; 1999.


32. Vatone H, Kortesito T, Kall M, for the EBMeDS Study Group. What may help or hinder the implementation of computerized decision support systems (CDSSs): a focus group study with physicians. Fam Pract 2008;25:162-7.


Patients with idiopathic/nonspecific low back pain of 4 weeks to 3 months duration without clinical concern for spine infection, malignancy, traumatic injury, or other serious conditions.

<table>
<thead>
<tr>
<th>Patient Symptoms Same or Better</th>
</tr>
</thead>
<tbody>
<tr>
<td>PHARMACOLOGIC</td>
</tr>
<tr>
<td>□ Acetaminophen</td>
</tr>
<tr>
<td>□ NSAID</td>
</tr>
<tr>
<td>□ Antidepressants (TCA*)</td>
</tr>
<tr>
<td>□ Benzodiazepines</td>
</tr>
<tr>
<td>□ Tramadol</td>
</tr>
<tr>
<td>□ Opioids</td>
</tr>
<tr>
<td>□ Other</td>
</tr>
<tr>
<td>ACTIVITY</td>
</tr>
<tr>
<td>□ Normal</td>
</tr>
<tr>
<td>□ Heat application</td>
</tr>
<tr>
<td>□ Massage</td>
</tr>
<tr>
<td>□ Yoga</td>
</tr>
<tr>
<td>□ Acupuncture</td>
</tr>
<tr>
<td>□ Progressive relaxation</td>
</tr>
<tr>
<td>□ Cognitive-behavioral therapy</td>
</tr>
<tr>
<td>□ Other</td>
</tr>
</tbody>
</table>

* TCA - tricyclic antidepressants

**Fig 2.** Nonspecific acute low back pain pathway: follow-up visit. NSAID = nonsteroidal anti-inflammatory drug.
Patients with idiopathic/nonspecific low back pain of 4 weeks to 3 months duration without clinical concern for spine infection, malignancy, traumatic injury, or other serious conditions.

**Patient symptoms worsened and/or signs/symptoms of spinal stenosis or radiculopathy**

<table>
<thead>
<tr>
<th>PHARMACOLOGIC</th>
<th>IMAGING</th>
</tr>
</thead>
<tbody>
<tr>
<td>☐ Acetaminophen</td>
<td>No previous spine surgery:</td>
</tr>
<tr>
<td>☐ NSAID</td>
<td>☐ Lumbar spine MRI without contrast (preferred)</td>
</tr>
<tr>
<td>☐ Antidepressants (TCA*)</td>
<td>☐ Lumbar spine CT without contrast (if MRI is nondiagnostic or unavailable)</td>
</tr>
<tr>
<td>☐ Benzodiazepines</td>
<td>☐ Other</td>
</tr>
<tr>
<td>☐ Tramadol</td>
<td>Previous lumbar spine surgery</td>
</tr>
<tr>
<td>☐ Opioids</td>
<td>☐ Lumbar spine MRI without and with contrast (preferred)</td>
</tr>
<tr>
<td>☐ Other</td>
<td>☐ Lumbar spine CT without contrast</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ACTIVITY</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>☐ Normal</td>
<td>☐ Other</td>
</tr>
<tr>
<td>☐ Heat application</td>
<td>CONSULTS</td>
</tr>
<tr>
<td>☐ Massage</td>
<td>☐ Physical Therapy</td>
</tr>
<tr>
<td>☐ Yoga</td>
<td>☐ Occupational Therapy</td>
</tr>
<tr>
<td>☐ Acupuncture</td>
<td>☐ Psychiatry</td>
</tr>
<tr>
<td>☐ Progressive relaxation</td>
<td>☐ Neurology</td>
</tr>
<tr>
<td>☐ Cognitive-behavioral therapy</td>
<td>☐ Neurosurgery</td>
</tr>
<tr>
<td>☐ Other</td>
<td>☐ Orthopedics</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>EDUCATION</th>
<th>RETURN APPOINTMENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>☐ Back Pain Pamphlet (eg, The Back Book†)</td>
<td>☐ 4 weeks</td>
</tr>
<tr>
<td>☐ Other</td>
<td>☐ Other</td>
</tr>
</tbody>
</table>

* TCA - tricyclic antidepressants

Fig 3. Nonspecific acute low back pain pathway: follow-up visit. NSAID = nonsteroidal anti-inflammatory drug.
Radiculopathy: Dysfunction of a nerve root associated with pain, sensory impairment, weakness, or diminished deep tendon reflexes in nerve root distribution.

Spinal Stenosis: Low back or radicular pain that increases with walking and improves with flexion (sitting or propping).

<table>
<thead>
<tr>
<th><strong>PHARMACOLOGIC</strong></th>
<th><strong>EDUCATION</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>□ Acetaminophen</td>
<td>□ Back Pain Pamphlet (eg, The Back Book†)</td>
</tr>
<tr>
<td>□ NSAID</td>
<td>□ Other</td>
</tr>
<tr>
<td>□ Antidepressants (TCA*)</td>
<td>□ CONSULTS</td>
</tr>
<tr>
<td>□ Benzodiazepines</td>
<td>□ Physical Therapy</td>
</tr>
<tr>
<td>□ Tramadol</td>
<td>□ Occupational Therapy</td>
</tr>
<tr>
<td>□ Opioids</td>
<td>□ Other</td>
</tr>
<tr>
<td>□ Gabapentin</td>
<td>□ RETURN APPOINTMENT</td>
</tr>
<tr>
<td>□ Other</td>
<td>□ 4 weeks</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>ACTIVITY</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>□ Normal</td>
</tr>
<tr>
<td>□ Heat application</td>
</tr>
<tr>
<td>□ Massage</td>
</tr>
<tr>
<td>□ Yoga</td>
</tr>
<tr>
<td>□ Acupuncture</td>
</tr>
<tr>
<td>□ Progressive relaxation</td>
</tr>
<tr>
<td>□ Cognitive-behavioral therapy</td>
</tr>
<tr>
<td>□ Other</td>
</tr>
</tbody>
</table>

* TCA - tricyclic antidepressants

Fig 5. Acute low back pain with symptoms of spinal stenosis or radiculopathy pathway: follow-up visit. NSAID = nonsteroidal anti-inflammatory drug.
Radicularopathy: Dysfunction of a nerve root associated with pain, sensory impairment, weakness, or diminished deep tendon reflexes in nerve root distribution.

Spinal Stenosis: Low back or radicular pain that increases with walking and improves with flexion (sitting or propping)

### Unchanged or Worsening signs/symptoms of Spinal Stenosis/Radiculopathy

<table>
<thead>
<tr>
<th>PHARMACOLOGIC</th>
<th>IMAGING (surgery and/or intervention candidate)</th>
</tr>
</thead>
<tbody>
<tr>
<td>□ Acetaminophen</td>
<td>No previous spine surgery</td>
</tr>
<tr>
<td>□ NSAID</td>
<td>□ Lumbar spine MRI without contrast (preferred)</td>
</tr>
<tr>
<td>□ Antidepressants (TCA*)</td>
<td>□ Lumbar spine CT without contrast (if MRI is nondiagnostic or unavailable)</td>
</tr>
<tr>
<td>□ Benzodiazepines</td>
<td>□ Other</td>
</tr>
<tr>
<td>□ Tramadol</td>
<td>Previous lumbar spine surgery</td>
</tr>
<tr>
<td>□ Opioids</td>
<td>□ Lumbar spine MRI without and with contrast (preferred)</td>
</tr>
<tr>
<td>□ Gabapentin</td>
<td>□ Lumbar spine CT without contrast</td>
</tr>
<tr>
<td>□ Other</td>
<td>□ Lumbar spine MRI without contrast (if contrast contraindication)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ACTIVITY</th>
<th>CONULTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>□ Normal</td>
<td>□ Other</td>
</tr>
<tr>
<td>□ Heat application</td>
<td>□ Physical Therapy</td>
</tr>
<tr>
<td>□ Massage</td>
<td>□ Occupational Therapy</td>
</tr>
<tr>
<td>□ Yoga</td>
<td>□ Psychiatry</td>
</tr>
<tr>
<td>□ Acupuncture</td>
<td>□ Neurology</td>
</tr>
<tr>
<td>□ Progressive relaxation</td>
<td>□ Neurosurgery</td>
</tr>
<tr>
<td>□ Cognitive-behavioral therapy</td>
<td>□ Orthopedics</td>
</tr>
<tr>
<td>□ Other</td>
<td>□ Other</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>EDUCATION</th>
<th>RETURN APPOINTMENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>□ Back Pain Pamphlet (eg, The Back Book†)</td>
<td>□ 4 weeks</td>
</tr>
<tr>
<td>□ Other</td>
<td>□ Other</td>
</tr>
</tbody>
</table>

* TCA - tricyclic antidepressants

Fig 6. Acute low back pain with symptoms of spinal stenosis or radiculopathy pathway: follow-up visit. NSAID = nonsteroidal anti-inflammatory drug.
Guideline Title

ACR Appropriateness Criteria® acute trauma to the knee.

Bibliographic Source(s)

Guideline Status
This is the current release of the guideline.


The appropriateness criteria are reviewed biennially and updated by the panels as needed, depending on introduction of new and highly significant scientific evidence.

Scope

Disease/Condition(s)
Acute trauma to the knee

Guideline Category
Diagnosis

Clinical Specialty
Emergency Medicine
Family Practice
Internal Medicine
Nuclear Medicine
Orthopedic Surgery
Radiology
Sports Medicine

Intended Users
Health Plans
Hospitals
Managed Care Organizations
Physicians
Utilization Management

Guideline Objective(s)
To evaluate the appropriateness of initial radiologic examinations for patients with acute trauma to the knee

Target Population
Patients with acute trauma to the knee
Interventions and Practices Considered

1. X-ray knee
2. Magnetic resonance imaging (MRI)
   • Knee without contrast
   • Knee without and with contrast
3. Ultrasound (US) knee
4. Magnetic resonance angiography (MRA)
   • Knee without and with contrast
   • Knee without contrast
5. Computed tomography (CT)
   • Knee without contrast
   • Knee with contrast
   • Knee without and with contrast
6. Technetium (Tc)-99m bone scan with single photon emission computed tomography (SPECT) lower extremity
7. Arteriography lower extremity
8. Computed tomography angiography (CTA) lower extremity with contrast

Major Outcomes Considered

Utility of radiologic examinations in differential diagnosis

Methodology

Methods Used to Collect/Select the Evidence

Searches of Electronic Databases

Description of Methods Used to Collect/Select the Evidence

Literature Search Procedure

The Medline literature search is based on keywords provided by the topic author. The two general classes of keywords are those related to the condition (e.g., ankle pain, fever) and those that describe the diagnostic or therapeutic intervention of interest (e.g., mammography, MRI).

The search terms and parameters are manipulated to produce the most relevant, current evidence to address the American College of Radiology Appropriateness Criteria (ACR AC) topic being reviewed or developed. Combining the clinical conditions and diagnostic modalities or therapeutic procedures narrows the search to be relevant to the topic. Exploding the term "diagnostic imaging" captures relevant results for diagnostic topics.

The following criteria/limits are used in the searches.

1. Articles that have abstracts available and are concerned with humans.
2. Restrict the search to the year prior to the last topic update or in some cases the author of the topic may specify which year range to use in the search. For new topics, the year range is restricted to the last 5 years unless the topic author provides other instructions.
3. May restrict the search to Adults only or Pediatrics only.
4. Articles consisting of only summaries or case reports are often excluded from final results.

The search strategy may be revised to improve the output as needed.

Number of Source Documents

The total number of source documents identified as the result of the literature search is not known.

Methods Used to Assess the Quality and Strength of the Evidence

Weighting According to a Rating Scheme (Scheme Given)

Rating Scheme for the Strength of the Evidence

Strength of Evidence Key

Category 1 - The conclusions of the study are valid and strongly supported by study design, analysis, and results.
Category 2 - The conclusions of the study are likely valid, but study design does not permit certainty.
Category 3 - The conclusions of the study may be valid, but the evidence supporting the conclusions is inconclusive or equivocal.
Category 4 - The conclusions of the study may not be valid because the evidence may not be reliable given the study design or analysis.
Methods Used to Analyze the Evidence

Systematic Review with Evidence Tables

Description of the Methods Used to Analyze the Evidence

The topic author drafts or revises the narrative text summarizing the evidence found in the literature. American College of Radiology (ACR) staff draft an evidence table based on the analysis of the selected literature. These tables rate the strength of the evidence for all articles included in the narrative text.

The expert panel reviews the narrative text, evidence table, and the supporting literature for each of the topic-variant combinations and assigns an appropriateness rating for each procedure listed in the table. Each individual panel member forms his/her own opinion based on his/her interpretation of the available evidence.

More information about the evidence table development process can be found in the ACR Appropriateness Criteria® Evidence Table Development document (see the “Availability of Companion Documents” field).

Methods Used to Formulate the Recommendations

Expert Consensus (Delphi)

Description of Methods Used to Formulate the Recommendations

Modified Delphi Technique

The appropriateness ratings for each of the procedures included in the Appropriate Criteria topics are determined using a modified Delphi methodology. A series of surveys are conducted to elicit each panelist’s expert interpretation of the evidence, based on the available data, regarding the appropriateness of an imaging or therapeutic procedure for a specific clinical scenario. American College of Radiology (ACR) staff distributes surveys to the panelists along with the evidence table and narrative. Each panelist interprets the available evidence and rates each procedure. The surveys are completed by panelists without consulting other panelists. The ratings are a scale between 1 and 9, which is further divided into three categories: 1, 2, or 3 is defined as “usually not appropriate”; 4, 5, or 6 is defined as “may be appropriate”; and 7, 8, or 9 is defined as “usually appropriate.” Each panel member assigns one rating for each procedure per survey round. The surveys are collected and the results are tabulated, de-identified and redistributed after each round. A maximum of three rounds are conducted. The modified Delphi technique enables each panelist to express individual interpretations of the evidence and his or her expert opinion without excessive bias from fellow panelists in a simple, standardized and economical process.

Consensus among the panel members must be achieved to determine the final rating for each procedure. Consensus is defined as eighty percent (80%) agreement within a rating category. The final rating is determined by the median of all the ratings once consensus has been reached. Up to three rating rounds are conducted to achieve consensus.

If consensus is not reached, the panel is convened by conference call. The strengths and weaknesses of each imaging procedure that has not reached consensus are discussed and a final rating is proposed. If the panelists on the call agree, the rating is accepted as the panel’s consensus. The document is circulated to all the panelists to make the final determination. If consensus cannot be reached on the call or when the document is circulated, “No consensus” appears in the rating column and the reasons for this decision are added to the comment sections.

Rating Scheme for the Strength of the Recommendations

Not applicable

Cost Analysis

The guideline developers reviewed published cost analyses.

Method of Guideline Validation

Internal Peer Review

Description of Method of Guideline Validation

Criteria developed by the Expert Panels are reviewed by the American College of Radiology (ACR) Committee on Appropriateness Criteria.

Recommendations

Major Recommendations

ACR Appropriateness Criteria®

Clinical Condition: Acute Trauma to the Knee

Variant 1: Patient any age (excluding infants); fall or twisting injury, no focal tenderness, no effusion; able to walk.

First study.

<table>
<thead>
<tr>
<th>Radiologic Procedure</th>
<th>Rating</th>
<th>Comments</th>
<th>RRL*</th>
</tr>
</thead>
<tbody>
<tr>
<td>X-ray knee</td>
<td>2</td>
<td></td>
<td>radioactive</td>
</tr>
<tr>
<td>MRI knee without contrast</td>
<td>2</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>MRI knee without and with contrast</td>
<td>1</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>US knee</td>
<td>1</td>
<td></td>
<td>0</td>
</tr>
</tbody>
</table>
### Exercise 1: Read and Analyze the Document

The document contains a table summarizing the radiation levels for various radiologic procedures. The table includes columns for the procedure, rating, comments, and relative radiation level (RRL) with a key for interpretation. The procedures listed include MRI knee, CT knee, and bone scans with SPECT. The ratings range from 1 (very appropriate) to 9 (very inappropriate), and the RRLs range from O (no radiation) to 3 (very high radiation). The comments section provides additional notes on the procedures.

#### Table: Relative Radiation Levels

<table>
<thead>
<tr>
<th>Radiologic Procedure</th>
<th>Rating</th>
<th>Comments</th>
<th>RRL*</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRI knee</td>
<td>9</td>
<td></td>
<td>radioactive</td>
</tr>
<tr>
<td>CT knee without contrast</td>
<td>1</td>
<td>The RRL for the adult procedure is radioactive</td>
<td>radioactive</td>
</tr>
<tr>
<td>CT knee with contrast</td>
<td>1</td>
<td>The RRL for the adult procedure is radioactive</td>
<td>radioactive</td>
</tr>
<tr>
<td>CT knee without and with contrast</td>
<td>1</td>
<td>The RRL for the adult procedure is radioactive</td>
<td>radioactive</td>
</tr>
<tr>
<td>Tc-99m bone scan with SPECT lower extremity</td>
<td>1</td>
<td></td>
<td>radioactive</td>
</tr>
</tbody>
</table>

#### Rating Scale: 1,2,3 Usually not appropriate; 4,5,6 May be appropriate; 7,8,9 Usually appropriate

**Note:** Abbreviations used in the tables are listed at the end of the "Major Recommendations" field.

**Variant 2:** Patient any age (excluding infants); fall or twisting injury, with one or more of following: focal tenderness, effusion, inability to bear weight. First study.

<table>
<thead>
<tr>
<th>Radiologic Procedure</th>
<th>Rating</th>
<th>Comments</th>
<th>RRL*</th>
</tr>
</thead>
<tbody>
<tr>
<td>X-ray knee</td>
<td>9</td>
<td></td>
<td>radioactive</td>
</tr>
<tr>
<td>MRI knee without contrast</td>
<td>5</td>
<td></td>
<td>O</td>
</tr>
<tr>
<td>US knee</td>
<td>2</td>
<td></td>
<td>O</td>
</tr>
<tr>
<td>CT knee without contrast</td>
<td>2</td>
<td>The RRL for the adult procedure is radioactive</td>
<td>radioactive</td>
</tr>
<tr>
<td>Tc-99m bone scan with SPECT lower extremity</td>
<td>2</td>
<td></td>
<td>radioactive</td>
</tr>
<tr>
<td>MRI knee without and with contrast</td>
<td>1</td>
<td></td>
<td>O</td>
</tr>
<tr>
<td>MRI knee without and with contrast</td>
<td>1</td>
<td></td>
<td>O</td>
</tr>
<tr>
<td>MRI knee without contrast</td>
<td>1</td>
<td></td>
<td>O</td>
</tr>
<tr>
<td>CT knee with contrast</td>
<td>1</td>
<td>The RRL for the adult procedure is radioactive</td>
<td>radioactive</td>
</tr>
<tr>
<td>CT knee without and with contrast</td>
<td>1</td>
<td>The RRL for the adult procedure is radioactive</td>
<td>radioactive</td>
</tr>
</tbody>
</table>

**Rating Scale:** 1,2,3 Usually not appropriate; 4,5,6 May be appropriate; 7,8,9 Usually appropriate

**Variant 3:** Patient any age (excluding infants); fall or twisting injury with either no fracture or a Segond fracture seen on a radiograph, with one or more of the following: focal tenderness, effusion, inability to bear weight. Next study.

<table>
<thead>
<tr>
<th>Radiologic Procedure</th>
<th>Rating</th>
<th>Comments</th>
<th>RRL*</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRI knee without contrast</td>
<td>9</td>
<td></td>
<td>O</td>
</tr>
<tr>
<td>CT knee without contrast</td>
<td>5</td>
<td>The RRL for the adult procedure is radioactive</td>
<td>radioactive</td>
</tr>
<tr>
<td>US knee</td>
<td>1</td>
<td></td>
<td>O</td>
</tr>
<tr>
<td>MRI knee without and with contrast</td>
<td>1</td>
<td></td>
<td>O</td>
</tr>
<tr>
<td>MRI knee without and with contrast</td>
<td>1</td>
<td></td>
<td>O</td>
</tr>
<tr>
<td>CT knee with contrast</td>
<td>1</td>
<td>The RRL for the adult procedure is radioactive</td>
<td>radioactive</td>
</tr>
<tr>
<td>CT knee without and with contrast</td>
<td>1</td>
<td>The RRL for the adult procedure is radioactive</td>
<td>radioactive</td>
</tr>
</tbody>
</table>

**Note:** Abbreviations used in the tables are listed at the end of the "Major Recommendations" field.
Note: Abbreviations used in the tables are listed at the end of the "Major Recommendations" field.

**Variant 4:** Patient any age (excluding infants); fall or twisting injury with a tibial plateau fracture on a radiograph, with one or more of the following: focal tenderness, effusion, inability to bear weight. Next study.

<table>
<thead>
<tr>
<th>Radiologic Procedure</th>
<th>Rating</th>
<th>Comments</th>
<th>RRL*</th>
</tr>
</thead>
<tbody>
<tr>
<td>CT knee without contrast</td>
<td>9</td>
<td>May be helpful for treatment planning or prognosis. The RRL for the adult procedure is *radioactive.</td>
<td>radioactive</td>
</tr>
<tr>
<td>MRI knee without contrast</td>
<td>7</td>
<td></td>
<td>radioactive</td>
</tr>
<tr>
<td>US knee</td>
<td>1</td>
<td></td>
<td>radioactive</td>
</tr>
<tr>
<td>MRI knee without and with contrast</td>
<td>1</td>
<td></td>
<td>radioactive</td>
</tr>
<tr>
<td>MRA knee without and with contrast</td>
<td>1</td>
<td></td>
<td>radioactive</td>
</tr>
<tr>
<td>MRA knee without contrast</td>
<td>1</td>
<td></td>
<td>radioactive</td>
</tr>
<tr>
<td>CT knee with contrast</td>
<td>1</td>
<td>The RRL for the adult procedure is *radioactive.</td>
<td>radioactive</td>
</tr>
<tr>
<td>CT knee without and with contrast</td>
<td>1</td>
<td>The RRL for the adult procedure is *radioactive.</td>
<td>radioactive</td>
</tr>
<tr>
<td>Tc-99m bone scan with SPECT lower extremity</td>
<td>1</td>
<td></td>
<td>radioactive</td>
</tr>
</tbody>
</table>

**Rating Scale:** 1,2,3 Usually not appropriate; 4,5,6 May be appropriate; 7,8,9 Usually appropriate

Note: Abbreviations used in the tables are listed at the end of the "Major Recommendations" field.

**Variant 5:** Patient any age (excluding infants). Injury to knee 2 days ago, mechanism unknown. Focal patellar tenderness, effusion, able to walk. First study.

<table>
<thead>
<tr>
<th>Radiologic Procedure</th>
<th>Rating</th>
<th>Comments</th>
<th>RRL*</th>
</tr>
</thead>
<tbody>
<tr>
<td>X-ray knee</td>
<td>9</td>
<td></td>
<td>radioactive</td>
</tr>
<tr>
<td>MRI knee without contrast</td>
<td>5</td>
<td></td>
<td>radioactive</td>
</tr>
<tr>
<td>US knee</td>
<td>2</td>
<td></td>
<td>radioactive</td>
</tr>
<tr>
<td>CT knee without contrast</td>
<td>2</td>
<td>The RRL for the adult procedure is *radioactive.</td>
<td>radioactive</td>
</tr>
<tr>
<td>Tc-99m bone scan with SPECT lower extremity</td>
<td>2</td>
<td></td>
<td>radioactive</td>
</tr>
<tr>
<td>MRI knee without and with contrast</td>
<td>1</td>
<td></td>
<td>radioactive</td>
</tr>
<tr>
<td>MRA knee without and with contrast</td>
<td>1</td>
<td></td>
<td>radioactive</td>
</tr>
<tr>
<td>MRA knee without contrast</td>
<td>1</td>
<td></td>
<td>radioactive</td>
</tr>
<tr>
<td>CT knee with contrast</td>
<td>1</td>
<td>The RRL for the adult procedure is *radioactive.</td>
<td>radioactive</td>
</tr>
<tr>
<td>CT knee without and with contrast</td>
<td>1</td>
<td>The RRL for the adult procedure is *radioactive.</td>
<td>radioactive</td>
</tr>
</tbody>
</table>

**Rating Scale:** 1,2,3 Usually not appropriate; 4,5,6 May be appropriate; 7,8,9 Usually appropriate

Note: Abbreviations used in the tables are listed at the end of the "Major Recommendations" field.

**Variant 6:** Patient any age (excluding infants). Significant trauma to knee from motor vehicle accident, suspect posterior knee dislocation. First study.

<table>
<thead>
<tr>
<th>Radiologic Procedure</th>
<th>Rating</th>
<th>Comments</th>
<th>RRL*</th>
</tr>
</thead>
</table>

*Relative Radiation Level*
Readers with questions regarding guideline content are directed to contact the guideline developer. The American College of Radiology (ACR) provided the funding and the resources for these ACR Appropriateness Criteria. The ACR Web site identifies and redistributes after each round. A maximum of 30 notes and comments have been identified and are listed at the end of the "Major Recommendations" field.

### Summary of Literature Review

A 2001 report stated that there are 1.3 million annual visits to United States emergency departments because of acute knee trauma, and over $1 billion is spent on radiographs of the knee. The knee radiograph is the most common radiograph performed for trauma in the emergency room and has the lowest yield for diagnosing clinically significant fractures. A retrospective review of 1,967 patients with acute knee injuries revealed that 74.1% of patients had radiographs and only 5.2% had fractures. One study concluded that radiographs obtained for acute trauma do not reliably depict all important injuries and that the findings in 25% of knee radiographs obtained for acute trauma do not correlate with the clinical findings.

### Radiography

A prospective survey of the judgment and attitudes of experienced clinicians in the use of knee radiography in 1,040 patients with acute knee injuries showed that, despite its inability to accurately predict the probability of fracture and to discriminate between fracture and nonfracture cases, radiographs were usually ordered. The proportion of patients referred for knee radiographs varied from 65.9% to 84.6%. According to the physicians, radiographs were ordered for the following reasons: 1) patients expected it and would otherwise be dissatisfied; 2) the physician lacked confidence in the clinical examination, or the orthopedic surgeon considered the radiograph routine; and 3) possible medicolegal repercussions. These reasons and the patient’s demand for imaging are recognized as the reasons that implementation of ordering guidelines was not overwhelming.

Caution should be used when relying on clinical examination for diagnosing certain knee injuries. One study reported that the correct diagnosis of bilateral quadriceps tendon rupture was established in only 61% (17/28) of cases by history and clinical examination alone. Another study reported that fractures missed on clinical examination included fractures of the patella, tibial spine, and fibular head.

Clinical decision rules for the acutely injured knee suggest that radiographic examination of the knee following acute injury can be eliminated in most instances by applying specific clinical guidelines. A prospective and retrospective study of 334 patients reported that patients between 12 and 50 years of age suffering a fall or blunt trauma and unable to ambulate or those who sustained multiple trauma should be radiographed. These authors reported 92% sensitivity and 79% specificity for identifying clinically significant fractures. Their study also reported that applying the clinical decision rules could reduce the number of radiographs taken in the emergency room by 78%.

A group of researchers concluded that a clinically significant fracture can be excluded in patients older than 18 years who can walk without limping or if there was a twisting injury to the knee and no joint effusion. If an effusion was present on physical examination, the odds of a fracture were 7.5 times greater. Using this clinical decision rule, the sensitivity for detecting a knee fracture was 100%, and specificity was sufficient to eliminate the need for 29% of knee radiographs ordered in the emergency room.
Another group applied a clinical decision rule (Ottawa Knee Rule) using parameters based on age, palpable tenderness, and function. Under the rule patients with acute knee pain and one or more of the following parameters should have a radiographic examination if they:

- Are 55 years of age or older
- Have palpable tenderness over the head of the fibula
- Have isolated patellar tenderness
- Cannot flex the knee to 90 degrees
- Cannot bear weight immediately following the injury, or
- Cannot walk in the emergency room (after taking four steps)

This rule was applied prospectively in 1,047 adults with acute knee injuries, and it was determined that its application would result in a 28% relative reduction in the number of radiographs ordered, a decrease from 68.6% to 49.4%

A later study was performed to validate the Ottawa Knee Rule, and prospective validation analyzing 1,096 patients found it to be 100% sensitive for identifying knee fractures. The decision rule was interpreted correctly 96% of the time, and when applied, the probability of missing a fracture was zero. The decision rule was 100% sensitive for identifying a fracture in patients older than 18 years who were not referred from another hospital, returned for reassessment, had a knee injury for seven days, or had isolated skin lesions. The potential relative reduction in use of radiography was estimated to be 28% (from 74% to 53%).

In a pooled analysis of data from six studies, it was concluded that a negative result using the Ottawa Knee Rule accurately excluded knee fracture after acute knee injury. A meta-analysis to determine the role of radiography in evaluating knee fractures concluded that among the five decision rules evaluated, the Ottawa Knee Rule had the strongest supporting evidence. Further prospective analysis of the Ottawa Knee Rule showed that it allowed a decrease in the number of radiographs performed after knee trauma by 35%, with a sensitivity of knee fracture detection of 100%.

Another study compared the implementation of the Ottawa Knee Rule by triage nurses and emergency medicine physicians. No fracture was missed by either group, but triage nurses were found to order 3.6 times more radiographs than emergency physicians, maintaining sensitivity at the expense of specificity and cost savings. An additional study evaluated the use of the Ottawa Knee Rule when applied by users with different levels of clinical training, including medical students and surgical residents, and found sensitivity and negative predictive value of 1.0 for both groups and a reduced radiography rate of 25% with application of the rule.

In a study of 214 patients, it was determined that the use of radiographs in the setting of acute trauma could be further reduced by obtaining a single lateral view. It was reported that the probability of not having a fracture if the lateral view was normal was 100%, thus reducing the need for additional radiographs by 67%.

With regard to mechanism of injury, history and physical examination are key elements for determining the indication for radiographs and the application of a decision rule. The most common mechanisms for knee injury are a direct blow, a fall, or a twisting injury. Twisting injuries are responsible for three-fourths of all knee injuries; however, 86% of all knee fractures result from blunt trauma. The risk of fracture also increases with age; fracture is four times more likely in patients older than 50 years, presumably secondary to osteoporosis, increased frequency of blunt injury, and inability to protect the knee during a fall.

Absence of immediate swelling, ecchymosis, effusion, deformity, increased warmth, and abrasion/laceration are significant predictors of a normal radiograph. It was generally agreed that radiographs should be obtained and that the clinical decision rule should not be applied for patients with gross deformity, a palpable mass, a penetrating injury, prosthetic hardware, unreliable clinical history or physical examination secondary to multiple injuries, altered mental status (e.g., head injury, drug or alcohol use, dementia), neuropathy (e.g., paraplegia, diabetes), or a history suggesting increased risk of fracture. The physician’s judgment and common sense, however, should supersede clinical guidelines.

Transient patellar dislocation is unsuspected clinically in 45% to 73% of patients with evidence of dislocation subsequently seen on magnetic resonance imaging (MRI). Radiographs may demonstrate a fracture of the medial patella or lateral trochlear, and can also show anatomic features that predispose to dislocation such as a decreased sulcus angle, patella alta, patellar tilt, or patellar subluxation. MRI is more sensitive than radiographs for detecting lateral patellar dislocation, including injury to the medial patellofemoral ligament, bone contusions and osteochondral injuries.

**Magnetic Resonance Imaging**

In addition to clinically significant fractures, other injuries must be considered. Most patients (93.5%) who present with acute knee injuries in the emergency room have soft-tissue rather than osseous injuries. Even in patients with fractures, concomitant soft-tissue injuries frequently are present. A study found that in 90% of patients with otherwise nonoperative tibial plateau fractures there were significant soft-tissue injuries diagnosed by MRI, including ligament and meniscal tears. Another study reported unstable meniscal tears in 36% of patients with tibial plateau fractures. An accurate clinical examination is essential to identify patients at high risk for delayed function recovery due to major soft-tissue injuries. However, using MRI, another study showed that the first clinical examination after acute knee trauma has a low diagnostic value and that the incidence of anterior cruciate ligament (ACL) injuries is higher than previously described. It is recognized that MRI is the optimal imaging modality for identifying soft-tissue, cartilaginous surface, and bone injuries around the knee.

To image internal knee derangement, MRI has been the technique of choice since the 1990s. The accuracy and reliability of MRI depend on experience and training. Nonetheless, numerous studies have shown that MRI has a high diagnostic accuracy in identifying traumatic intra-articular knee lesions. This is particularly true when strict diagnostic criteria are used, and this applies to both spin-echo imaging and fast spin-echo imaging, as well as imaging at both low and high field strength. MRI has been shown to demonstrate minor meniscocapsular tears when performed with understanding of anatomy. Characteristic findings on MRI, including specific bone marrow edema patterns and osteochondral defects, can allow accurate diagnosis of injuries such as transient dislocation of the patella that cannot be detected by radiographs.

MRI is a valuable tool in the decision-making process, altering the treatment plan in 18% of patients with meniscal or chondral surface injuries and allowing earlier surgical intervention because of the more accurate diagnosis obtained. Multiple authors and studies have validated that unnecessary diagnostic arthroscopy can be avoided because of the high predictive value of a negative MRI. One study found MRI to have a positive predictive value twice that of clinical examination for meniscal tears. It also found that MRI would decrease negative diagnostic arthroscopy to 5% and would help reduce the need for a second therapeutic arthroscopic procedure. Another study reported MRI accuracy to be...
ACL rupture is responsible for more than 70% of all acute hemarthrosis in young athletes and 17% in a mixed sedentary and athletic population. Locking, the presence of a loose body on radiographs, and hemarthrosis within 12 hours of injury have previously been reported as indications for arthroscopy instead of MRI. However, it was found that in 48% of patients presenting with an acutely locked knee, management was changed from surgical to conservative based on MRI findings.

**Single Photon Emission Computed Tomography**

In addition to MRI, single photon emission computed tomography (SPECT) has been proposed for diagnosing meniscus injuries. A specific crescentic pattern of uptake on the transaxial view has been described as having a sensitivity of 77% and specificity of 74%. With the additional criterion of increased equilibrium activity in the adjacent femoral condyles, these values increase to 90% and 84%, respectively. Considerable concordance has been shown between SPECT results and those of other modalities for assessing meniscal tears and the bone contusions from an ACL tear in acute knee trauma.

**Ultrasound**

Sonography has been reported to be 91% sensitive and 100% specific for diagnosing an acute ACL tear within 10 weeks of an acute hemarthrosis when there is no prior trauma and no bone abnormalities. Sonography can be used both for initial detection and confirmation of this injury and for follow-up. Furthermore, a comparison of sonography and arthrography using lipohemarthrosis as a criterion of osteoarticular fracture has revealed a sensitivity and specificity of 94% for sonographic detection of such fractures. One study showed that the presence of an effusion at sonography in the acutely injured knee has a 91% positive predictive value for internal derangement. However, intra-articular knee sonography should only be performed and interpreted by personnel with the appropriate expertise in its application.

**Computed Tomography**

Computed tomography (CT) with three-dimensional reconstruction has been shown to reflect the severity of tibial plateau fractures more accurately than radiography in 43% of cases and to modify the surgical plan in 59% of operative cases. In severely injured patients, diagnostically sufficient radiographs are sometimes difficult to obtain, and therefore a negative radiograph is not reliable in ruling out a fracture. In these patients, multidetector CT is a fast and accurate examination for evaluating tibial plateau fractures and other complex knee injuries. In a 2007 study, researchers concluded that in the acute setting, CT offers 80% sensitivity and 98% specificity for depicting osseous avulsions and a high negative predictive value for excluding ligament injury.

**Patellar Dislocation**

Transient patellar dislocation is unsuspected clinically in 45% to 73% of patients with evidence of dislocation subsequently seen on MRI. Radiographs may demonstrate a fracture of the medial patella or lateral trochlear, and can also show anatomic features that predispose to dislocation such as a decreased sulcus angle, patella alta, patellar tilt, or patellar subluxation. MRI is more sensitive than radiographs for imaging findings of lateral patellar dislocation, including injury to the medial patellofemoral ligament, bone contusions, and osteochondral injuries.

**Knee Dislocation**

Dislocation of the knee results from a fall from a height, a motor vehicle accident, a vehicle striking a pedestrian, or contact sports. This injury, which often reduces spontaneously, constitutes a true orthopedic emergency because of possible nerve or arterial damage. Vascular injury may be found in one-third of patients following posterior knee dislocation. Physical signs of clinically significant vascular injury are the absence of pulses, ischemia, active bleeding, and bruit. These signs have been reported to have 100% intra-articular fracture as a criterion of osteoarticular fracture was established in only 61% (17/28) of cases by history and ordering guidelines was not overwhelming.

**Summary**

Clinical decision rules for evaluating the acutely injured knee have been studied by various investigators, who determined that their application can considerably reduce the number of radiographs ordered without missing a clinically significant fracture. Although different parameters and definitions were used for the various decision rules, there were sufficient similarities between the investigations to allow usable conclusions to be drawn.

In patients of any age except for infants, the clinical parameters used for not requiring radiographs following knee trauma are as follows:

- Patient is able to walk without a limp
- Patient had a twisting injury and there is no effusion

The clinical parameters for ordering knee radiographs in this population following trauma are as follows:

- Joint effusion within 24 hours of direct blow or fall
- Palpable tenderness over the fibular head or patella
Inability to walk (four steps) or bear weight immediately or in the emergency room or within a week of the trauma.

Inability to flex knee to 90 degrees

Altered mental status

It has also been reported that a fracture can be excluded if a single lateral view of the knee is normal, eliminating the need for additional radiographic views.

In general, these studies excluded patients with superficial skin injuries, gross deformity, a palpable mass, a penetrating injury, prosthetic hardware, altered consciousness (from alcohol and/or drug use), multiple injuries, decreased limb sensation, or a history indicating an elevated risk of fracture. They also excluded pregnant patients, patients returning for reassessment, and patients whose injury occurred more than 7 days prior to initial evaluation.

Soft-tissue injuries (meniscal injuries, chondral surface injuries, and ligamentous disruption) are best evaluated by MRI. Although lateral patellar dislocation may be reduced at the time of presentation in the emergency room, characteristic findings on MRI, including specific bone marrow edema patterns and osteochondral defects, can allow accurate diagnosis. Knee dislocation, even if spontaneously reduced, constitutes a potential threat to the popliteal nerve or artery. A systematic review has suggested that the isolated presence of abnormal pedal pulses on initial examination following knee dislocation is not sensitive enough to detect a vascular injury that necessitates surgery, and that the workup should include angiography. One study has shown a 100% correlation between MRA findings and conventional angiography findings in multiple-ligament injured knees, including knee dislocations. An MRI should also be performed to identify ligamentous injuries and associated pathology.

Abbreviations

- CT, computed tomography
- CTA, computed tomography angiography
- MRA, magnetic resonance angiography
- MRI, magnetic resonance imaging
- SPECT, single photon emission computed tomography
- Tc-99m, technetium-99 metastable
- US, ultrasound

Relative Radiation Level Designations

<table>
<thead>
<tr>
<th>Relative Radiation Level*</th>
<th>Adult Effective Dose Estimate Range</th>
<th>Pediatric Effective Dose Estimate Range</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0 mSv</td>
<td>0 mSv</td>
</tr>
<tr>
<td>radioactive</td>
<td>&lt;0.1 mSv</td>
<td>&lt;0.03 mSv</td>
</tr>
<tr>
<td>radioactive</td>
<td>0.1-1 mSv</td>
<td>0.03-0.3 mSv</td>
</tr>
<tr>
<td>radioactive</td>
<td>1-10 mSv</td>
<td>0.3-3 mSv</td>
</tr>
<tr>
<td>radioactive</td>
<td>10-30 mSv</td>
<td>3-10 mSv</td>
</tr>
<tr>
<td>radioactive</td>
<td>30-100 mSv</td>
<td>10-30 mSv</td>
</tr>
<tr>
<td>radioactive</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

PRRL assignments for some of the examinations cannot be made, because the actual patient doses in these procedures vary as a function of a number of factors (e.g., region of the body exposed to ionizing radiation, the imaging guidance that is used). The RRIs for these examinations are designated as “Varies”.

Clinical Algorithm(s)

Algorithms were not developed from criteria guidelines.

Evidence Supporting the Recommendations

Type of Evidence Supporting the Recommendations

The recommendations are based on analysis of the current literature and expert panel consensus.

Benefits/Harms of Implementing the Guideline Recommendations

Potential Benefits
Selection of appropriate radiologic imaging procedures for evaluation of patients with acute trauma to the knee

Potential Harms

Relative Radiation Level (RRL)

Potential adverse health effects associated with radiation exposure are an important factor to consider when selecting the appropriate imaging procedure. Because there is a wide range of radiation exposures associated with different diagnostic procedures, an RRL indication has been included for each imaging examination. The RRLs are based on effective dose, which is a radiation dose quantity that is used to estimate population total radiation risk associated with an imaging procedure. Patients in the pediatric age group are at inherently higher risk from exposure, both because of organ sensitivity and longer life expectancy (relevant to the long latency that appears to accompany radiation exposure). For these reasons, the RRL dose estimate ranges for pediatric examinations are lower as compared to those specified for adults. Additional information regarding radiation dose assessment for imaging examinations can be found in the American College of Radiology (ACR) Appropriateness Criteria® Radiation Dose Assessment Introduction document (see the “Availability of Companion Documents” field).

Qualifying Statements

The American College of Radiology (ACR) Committee on Appropriateness Criteria and its expert panels have developed criteria for determining appropriate imaging examinations for diagnosis and treatment of specified medical condition(s). These criteria are intended to guide radiologists, radiation oncologists, and referring physicians in making decisions regarding radiologic imaging and treatment. Generally, the complexity and severity of a patient’s clinical condition should dictate the selection of appropriate imaging procedures or treatments. Only those examinations generally used for evaluation of the patient’s condition are ranked. Other imaging studies necessary to evaluate other co-existent diseases or other medical consequences of this condition are not considered in this document. The availability of equipment or personnel may influence the selection of appropriate imaging procedures or treatments. Imaging techniques classified as investigational by the U.S. Food and Drug Administration (FDA) have not been considered in developing these criteria; however, study of new equipment and applications should be encouraged. The ultimate decision regarding the appropriateness of any specific radiologic examination or treatment must be made by the referring physician and radiologist in light of all the circumstances presented in an individual examination.

Implementation of the Guideline

Description of Implementation Strategy

An implementation strategy was not provided.

IOM Care Need

Getting Better

IOM Domain

Effectiveness

Identifying Information and Availability

Bibliographic Source(s)


Adaptation

Not applicable: The guideline was not adapted from another source.

Date Released

1998 (revised 2011)

Guideline Developer(s)

American College of Radiology - Medical Specialty Society

Source(s) of Funding

The American College of Radiology (ACR) provided the funding and the resources for these ACR Appropriateness Criteria®.

Guideline Committee
Knee dislocation, even if spontaneously reduced, constitutes a potential threat to the popliteal nerve or artery. A comparison between MRI findings and surgical findings in patients with knee dislocation. Furthermore, these authors reported 100% accuracy of MRI in identifying associated vascular injuries. In the acute setting, CT offers 80% sensitivity and 98% specificity for depicting osseous avulsions and a high negative predictive value. The false-negative rate of CT for detecting vascular injuries in patients with knee dislocation is less than 10%. Multiple authors and studies have validated that unnecessary diagnostic arthroscopy can be avoided because of the high sensitivity of MRI of the knee.

In a study of 214 patients, it was determined that the use of radiographs in the setting of acute trauma could be further restricted. The vast majority (99%) of patients with knee dislocation had normal radiographs, and these results were not altered by consultation with an emergency physician. Performing knee radiographs in this patient population was shown to save on average 8.7 dollars, with a range of $6.59-$11.96. Additionally, an additional study demonstrated that performing radiographs did not change the clinical diagnosis of acute injury. A comparison of clinical findings with MRI in 44 patients with suspected acute knee injuries demonstrated that the correct diagnosis of quadriceps tendon rupture was established in only 61% (17/28) of cases by history and physical examination alone. The correct diagnosis was established in 93% (21/23) of cases in which MRI was ordered. Furthermore, in a retrospective study of 26 patients with knee dislocation, it was determined that the use of radiographs in the setting of acute trauma could be further restricted. The authors demonstrated that performing radiographs did not change the clinical diagnosis of acute injury.

Note: Abbreviations used in the table are listed at the end of the "Major Recommendations" field.

<table>
<thead>
<tr>
<th>Radiologic Procedure</th>
<th>Level</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>CT knee without contrast</td>
<td>1</td>
<td>O varying cost-effectiveness of the procedure based on age, body habitus, and other factors.</td>
</tr>
<tr>
<td>MRI knee without contrast</td>
<td>1</td>
<td>O varying cost-effectiveness of the procedure based on age, body habitus, and other factors.</td>
</tr>
<tr>
<td>MRA knee without and with contrast</td>
<td>1</td>
<td>O varying cost-effectiveness of the procedure based on age, body habitus, and other factors.</td>
</tr>
<tr>
<td>99m bone scan with SPECT lower extremity</td>
<td>1</td>
<td>O varies widely depending on the dose administered and the type of tracer used.</td>
</tr>
<tr>
<td>99m bone scan with SPECT upper extremity</td>
<td>1</td>
<td>O varies widely depending on the dose administered and the type of tracer used.</td>
</tr>
</tbody>
</table>

Evaluators determine the level of cost-effectiveness of the procedure based on age, body habitus, and other factors. The actual patient doses in these procedures vary as a function of a number of factors, including age, body habitus, and other factors.

Patient Resources

None available

NGC Status

This summary was completed by ECRI on May 6, 2001. The information was verified by the guideline developer as of June 29, 2001. This summary was updated by ECRI on July 31, 2002. The updated information was verified by the guideline developer on October 1, 2002. This summary was updated by ECRI on February 6, 2006. This summary was updated by ECRI Institute on May 18, 2010. This summary was updated by ECRI Institute on August 24, 2011.

Copyright Statement

Instructions for downloading, use, and reproduction of the American College of Radiology (ACR) Appropriateness Criteria® may be found on the ACR Web site.

Disclaimer

The National Guideline Clearinghouse™ (NGC) does not develop, produce, approve, or endorse the guidelines represented on this site.

All guidelines summarized by NGC and hosted on our site are produced under the auspices of medical specialty societies, relevant professional associations, public or private organizations, other government agencies, health care organizations or plans, and similar entities.

Guidelines represented on the NGC Web site are submitted by guideline developers, and are screened solely to determine that they meet the NGC Inclusion Criteria which may be found at http://www.guideline.gov/about/inclusion-criteria.aspx.
Panel Members
The American College of Radiology (ACR) provided the funding and the resources for these ACR Appropriateness Criteria. ACR Appropriateness Criteria (ACR) staff draft an evidence table based on the analysis of the selected literature. These tables rate the strength of the evidence for all articles included in the narrative text.

**Algorithm**

- **MRA knee without contrast**
- **MRA knee without and with contrast**
- **MRI knee without and with contrast**
- **CT knee without and with contrast**

**Level 1**

- **99m bone scan with SPECT lower limb**
- **Bone scan under local anesthesia**

**Level 2**

- **Ultrasound**
- **Positron emission tomography (PET)**
- **Technetium 99m bone scan (radionuclide bone scan)**
- **Stress radiographs**
- **Osteodensitometry**
- **Single photon emission computerized tomography (SPECT)**
- **Sulphur 35**

**Level 3**

- **Sulphur 35**
- **Osteodensitometry**

**Level 4**

- **Sulphur 35**

**Comments**

- **99m bone scan with SPECT lower limb**: This imaging modality is reported to have a sensitivity of 90% and a specificity of 94% for detecting bone injuries. It is particularly useful for evaluating lesions in non-weight-bearing bones, where other imaging techniques may be less effective.

- **Bone scan under local anesthesia**: This technique allows for more localized evaluation of bone lesions, which can be particularly useful in identifying small or superficial lesions.

**Rating Scale**

- **1**: Usually inappropriate
- **2**: May be appropriate
- **3**: Usually appropriate

**RRL**

- **RRL***: Radiation Dose

**Radiation Dose**

- **0 mSv**: Lower limb bone scan

**Reference**

ACR Appropriateness Criteria

**American College of Radiology (ACR); 2008. 8 p. [60 references]**

**This is the current release of the guideline.**
Guideline Summary NGC-7001

Guideline Title
ACR Appropriateness Criteria® nontraumatic knee pain.

Bibliographic Source(s)

Guideline Status
This is the current release of the guideline.
The appropriateness criteria are reviewed biennially and updated by the panels as needed, depending on introduction of new and highly significant scientific evidence.

Scope

Disease/Condition(s)
Nontraumatic knee pain

Guideline Category
Diagnosis
Evaluation

Clinical Specialty
Family Practice
Internal Medicine
Nuclear Medicine
Orthopedic Surgery
Pediatrics
Radiology

Intended Users
Health Plans
Hospitals
Managed Care Organizations
Physicians
Utilization Management

Guideline Objective(s)
To evaluate the appropriateness of initial radiologic examinations for patients with nontraumatic knee pain

Target Population
Patients with nontraumatic knee pain
Interventions and Practices Considered
1. X-ray, knee and ipsilateral hip
2. Magnetic resonance imaging (MRI), knee, without contrast
3. MR arthrography, knee
4. Nuclear medicine(NUC), - technetium (Tc)-99m bone scan, lower extremity
5. Ultrasound (US), knee
6. Computed tomography (CT), knee, without contrast
7. CT arthrography, knee

Major Outcomes Considered
Utility of radiologic examinations in differential diagnosis

Methodology

Methods Used to Collect/Select the Evidence
Searches of Electronic Databases

Description of Methods Used to Collect/Select the Evidence
The guideline developer performed literature searches of peer-reviewed medical journals, and the major applicable articles were identified and collected.

Number of Source Documents
Not stated

Methods Used to Assess the Quality and Strength of the Evidence
Weighting According to a Rating Scheme (Scheme Not Given)

Rating Scheme for the Strength of the Evidence
Not stated

Methods Used to Analyze the Evidence
Systematic Review with Evidence Tables

Description of the Methods Used to Analyze the Evidence
One or two topic leaders within a panel assume the responsibility of developing an evidence table for each clinical condition, based on analysis of the current literature. These tables serve as a basis for developing a narrative specific to each clinical condition.

Methods Used to Formulate the Recommendations
Expert Consensus (Delphi)

Description of Methods Used to Formulate the Recommendations
Since data available from existing scientific studies are usually insufficient for meta-analysis, broad-based consensus techniques are needed for reaching agreement in the formulation of the appropriateness criteria. The American College of Radiology (ACR) Appropriateness Criteria panels use a modified Delphi technique to arrive at consensus. Serial surveys are conducted by distributing questionnaires to consolidate expert opinions within each panel. These questionnaires are distributed to the participants along with the evidence table and narrative as developed by the topic leader(s). Questionnaires are completed by the participants in their own professional setting without influence of the other members. Voting is conducted using a scoring system from 1 to 9, indicating the least to the most appropriate imaging examination or therapeutic procedure. The survey results are collected, tabulated in anonymous fashion, and redistributed after each round. A maximum of three rounds is conducted and opinions are unified to the highest degree possible. Eighty percent agreement is considered a consensus. This modified Delphi technique enables individual, unbiased expression, is economical, easy to understand, and relatively simple to conduct.

If consensus cannot be reached by this Delphi technique, the panel is convened and group consensus techniques are utilized. The strengths and weaknesses of each test or procedure are discussed and consensus reached whenever possible. If "No consensus" appears in the rating column, reasons for this decision are added to the comment sections.

Rating Scheme for the Strength of the Recommendations
Not applicable

Cost Analysis
A formal cost analysis was not performed and published cost analyses were not reviewed.

**Method of Guideline Validation**

Internal Peer Review

**Description of Method of Guideline Validation**

Criteria developed by the Expert Panels are reviewed by the American College of Radiology (ACR) Committee on Appropriateness Criteria.

**Recommendations**

**Major Recommendations**

**Note from the American College of Radiology (ACR) and the National Guideline Clearinghouse (NGC):** ACR has updated its Relative Radiation Level categories and Rating Scale. The Rating Scale now includes categories (1, 2, 3 = Usually not appropriate; 4, 5, 6 = May be appropriate; 7, 8, 9 = Usually appropriate). See the original guideline document for details.

**ACR Appropriateness Criteria®**

**Clinical Condition: Nontraumatic Knee Pain**

**Variant 1:** Child or adolescent: nonpatellofemoral symptoms. Mandatory minimal initial exam.

<table>
<thead>
<tr>
<th>Radiologic Procedure</th>
<th>Rating</th>
<th>Comments</th>
<th>RRL*</th>
</tr>
</thead>
<tbody>
<tr>
<td>X-ray knee</td>
<td>9</td>
<td>Min</td>
<td></td>
</tr>
<tr>
<td>X-ray hip ipsilateral</td>
<td>1</td>
<td>Med</td>
<td></td>
</tr>
<tr>
<td>CT knee without contrast</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CT arthrography knee</td>
<td>1</td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td>MRI knee without contrast</td>
<td>1</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>MR arthrography knee</td>
<td>1</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>US knee</td>
<td>1</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>NUC Tc-99m bone scan lower extremity</td>
<td>1</td>
<td>Med</td>
<td></td>
</tr>
</tbody>
</table>

**Rating Scale:** 1=Least appropriate, 9=Most appropriate

*Relative Radiation Level

Note: Abbreviations used in the tables are listed at the end of the "Major Recommendations" field.

**Variant 2:** Child or adult: patellofemoral (anterior) symptoms. Mandatory minimal initial exam.

<table>
<thead>
<tr>
<th>Radiologic Procedure</th>
<th>Rating</th>
<th>Comments</th>
<th>RRL*</th>
</tr>
</thead>
<tbody>
<tr>
<td>X-ray knee</td>
<td>9</td>
<td>Min</td>
<td></td>
</tr>
<tr>
<td>X-ray hip ipsilateral</td>
<td>1</td>
<td>Med</td>
<td></td>
</tr>
<tr>
<td>CT knee without contrast</td>
<td>1</td>
<td>Min</td>
<td></td>
</tr>
<tr>
<td>CT arthrography knee</td>
<td>1</td>
<td>Min</td>
<td></td>
</tr>
<tr>
<td>MRI knee without contrast</td>
<td>1</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>MR arthrography knee</td>
<td>1</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>US knee</td>
<td>1</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>NUC Tc-99m bone scan lower extremity</td>
<td>1</td>
<td>Med</td>
<td></td>
</tr>
</tbody>
</table>

**Rating Scale:** 1=Least appropriate, 9=Most appropriate

*Relative Radiation Level

Note: Abbreviations used in the tables are listed at the end of the "Major Recommendations" field.

**Variant 3:** Adult: nontrauma, nontumor, nonlocalized pain. Mandatory minimal initial exam.

<table>
<thead>
<tr>
<th>Radiologic Procedure</th>
<th>Rating</th>
<th>Comments</th>
<th>RRL*</th>
</tr>
</thead>
<tbody>
<tr>
<td>X-ray knee</td>
<td>9</td>
<td>Min</td>
<td></td>
</tr>
<tr>
<td>X-ray hip ipsilateral</td>
<td>1</td>
<td>Med</td>
<td></td>
</tr>
<tr>
<td>CT knee without contrast</td>
<td>1</td>
<td>Min</td>
<td></td>
</tr>
<tr>
<td>CT arthrography knee</td>
<td>1</td>
<td>Min</td>
<td></td>
</tr>
<tr>
<td>MRI knee without contrast</td>
<td>1</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>MR arthrography knee</td>
<td>1</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>US knee</td>
<td>1</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>NUC Tc-99m bone scan lower extremity</td>
<td>1</td>
<td>Med</td>
<td></td>
</tr>
</tbody>
</table>

**Rating Scale:** 1=Least appropriate, 9=Most appropriate

*Relative Radiation Level

Note: Abbreviations used in the tables are listed at the end of the "Major Recommendations" field.

**Variant 4:** Child or adolescent: nonpatellofemoral symptoms. Initial knee radiographs nondiagnostic (demonstrate normal findings or a joint effusion).

<table>
<thead>
<tr>
<th>Radiologic Procedure</th>
<th>Rating</th>
<th>Comments</th>
<th>RRL*</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRI knee without contrast</td>
<td>9</td>
<td>None</td>
<td></td>
</tr>
</tbody>
</table>

Note: Abbreviations used in the tables are listed at the end of the "Major Recommendations" field.
Intraoperative radiographs in the knee are usually not indicated in patients with conventional evidence of inflammatory arthritis of the knee. The consensus of the panel is that a knee radiograph may be indicated if an additional injury is suspected clinically or when it is necessary to determine the status of the femoral condyle or of the medial tibial plateau associated with tibial stress fracture.

X-ray knee

CT knee without contrast
CT arthrography knee
MR arthrography knee
US knee
NUC Tc-99m bone scan lower extremity

**Rating Scale:** 1 = Least appropriate, 9 = Most appropriate

*Relative Radiation Level*
### Variant 9: Adult: nontumor, nonlocalized pain. Initial knee radiographs demonstrate inflammatory, crystalline, or degenerative joint disease (uni- to tri- compartmental sclerosis, hypertrophic spurs, joint space narrowing, and/or subchondral cysts).

<table>
<thead>
<tr>
<th>Radiologic Procedure</th>
<th>Rating</th>
<th>Comments</th>
<th>Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>X-ray hip ipsilateral</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CT knee without contrast</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CT arthrography knee</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MRI knee without contrast</td>
<td>1</td>
<td>Consider for preoperative assessment.</td>
<td>None</td>
</tr>
<tr>
<td>MRI arthrography knee</td>
<td>1</td>
<td></td>
<td>None</td>
</tr>
<tr>
<td>US knee</td>
<td>1</td>
<td></td>
<td>None</td>
</tr>
<tr>
<td>NUC Tc-99m bone scan lower extremity</td>
<td>1</td>
<td></td>
<td>Med</td>
</tr>
</tbody>
</table>

**Rating Scale:** 1=Least appropriate, 9=Most appropriate

*Relative Radiation Level*

### Variant 10: Adult: nontumor, nonlocalized pain. Initial knee radiographs demonstrate avascular necrosis.

<table>
<thead>
<tr>
<th>Radiologic Procedure</th>
<th>Rating</th>
<th>Comments</th>
<th>Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRI knee without contrast</td>
<td>7</td>
<td>If needed for therapy</td>
<td>None</td>
</tr>
<tr>
<td>CT knee without contrast</td>
<td>1</td>
<td></td>
<td>Min</td>
</tr>
<tr>
<td>CT arthrography knee</td>
<td>1</td>
<td></td>
<td>Min</td>
</tr>
<tr>
<td>MR arthrography knee</td>
<td>1</td>
<td></td>
<td>None</td>
</tr>
<tr>
<td>US knee</td>
<td>1</td>
<td></td>
<td>None</td>
</tr>
<tr>
<td>NUC Tc-99m bone scan lower extremity</td>
<td>1</td>
<td></td>
<td>Med</td>
</tr>
</tbody>
</table>

**Rating Scale:** 1=Least appropriate, 9=Most appropriate

*Relative Radiation Level*

### Variant 11: Adult: nontumor, nonlocalized pain. Initial knee radiographs demonstrate evidence of internal derangement (e.g., Segond fracture, deep lateral femoral notch sign).

<table>
<thead>
<tr>
<th>Radiologic Procedure</th>
<th>Rating</th>
<th>Comments</th>
<th>Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRI knee without contrast</td>
<td>9</td>
<td></td>
<td>None</td>
</tr>
<tr>
<td>CT arthrography knee</td>
<td>2</td>
<td>If MRI contraindicated.</td>
<td>Min</td>
</tr>
<tr>
<td>CT knee without contrast</td>
<td>1</td>
<td></td>
<td>Min</td>
</tr>
<tr>
<td>MR arthrography knee</td>
<td>1</td>
<td></td>
<td>None</td>
</tr>
<tr>
<td>US knee</td>
<td>1</td>
<td></td>
<td>None</td>
</tr>
<tr>
<td>NUC Tc-99m bone scan lower extremity</td>
<td>1</td>
<td></td>
<td>Med</td>
</tr>
</tbody>
</table>

**Rating Scale:** 1=Least appropriate, 9=Most appropriate

*Relative Radiation Level*

---

**Summary of Literature Review**

Nontraumatic knee pain in children, adolescents, and adults includes localized complaints such as anterior (patellofemoral) pain and diffuse nonlocalized symptoms. The consensus of the committee is that the initial imaging studies for nontraumatic knee pain are anteroposterior (AP) and lateral radiograph. For patients with diffuse non-localized symptoms, a Merchant or axial view may be useful as part of the initial examination. In children with nontraumatic knee pain, referred pain from the hip must be considered and hip radiographs may need to be obtained if there is clinical evidence or clinical concern for hip pathology.

In elderly patients, the most common source of nontraumatic knee pain is osteoarthritis. Conventional radiographic diagnosis of degenerative joint disease (osteoarthritis) includes joint space narrowing, osteophytes, subchondral cysts, and sclerosis bordering the joint. Articular cartilage is evaluated indirectly on radiographs by joint space narrowing and changes in the subchondral bone. Routine radiographs are insensitive for assessing articular cartilage in the early stages of osteoarthritis, while in advanced disease, joint space narrowing on radiographs is usually an accurate assessment of cartilage loss. Standing radiographs have been reported to more accurately reflect medial and lateral joint compartment cartilage loss than supine radiographs; however, in the presence of a severe varus or valgus deformity, significant cartilage loss in the compartment that appears wide (due to the alignment deformity) may not be evident. A weight-bearing posteroanterior (PA) radiograph, obtained with knee flexion, has been reported to show the cartilage width of the posterior medial and lateral joint compartments more accurately than that a standing view obtained with the knee extended. The standing flexed view may be indicated in elderly patients with osteoarthritis when surgical intervention is being planned. Finally, one should bear in mind that a significant portion of the joint space narrowing may be due to meniscal extrusion or degeneration rather than hyaline cartilage loss in some patients. Additional imaging studies are not indicated in patients for whom radiographs are diagnostic of degenerative joint disease unless treatment, or surgery, or both are dependent on additional findings such as internal knee derangement or when symptoms are not explained by the radiographic findings (e.g., stress fractures).

Other nontraumatic causes of knee pain in adult patients include internal knee derangement (meniscal and ligament tears), stress fracture, subchondral insufficiency fracture (known as spontaneous osteonecrosis), inflammatory arthritis, transient osteoporosis, and chronic regional pain syndrome. Chronic anterior lateral knee pain may also result from patella tendon–lateral femoral condyle friction syndrome or iliobibial band syndrome (friction syndrome) both of which can...
be confirmed or excluded by magnetic resonance imaging (MRI).

When initial radiographs are nondiagnostic (normal findings or a joint effusion) and knee symptoms are suspicious for an internal derangement, the next indicated study is an MRI examination. MRI is also indicated when the patient has persistent knee pain and normal radiographs. MRI is more sensitive than radiography and provides more specific information compared with radionuclide bone scan. MRI of nontraumatic knee pain may document a joint effusion, communicating synovial cysts, proliferative changes of the synovial membrane, osteophytes, subchondral cysts, articular cartilage loss, meniscal and/or ligamentous tears and/or degeneration, bone marrow edema, fractures, and osteonecrosis. A secondary MRI finding with a high sensitivity for internal derangement is an AP joint fluid measurement of greater than 10 mm in the lateral suprapatellar pouch.

MRI is useful to identify a subchondral insufficiency fracture as the initial injury from which localized osteonecrosis may result and which was otherwise identified as spontaneous osteonecrosis. MRI can also detect osteonecrosis of the medial femoral condyle or of the medial tibial plateau associated with tibial stress fracture.

A suprapatellar joint effusion is readily detected on a lateral radiograph of the knee; however, the extent of a joint effusion, the presence of a communicating synovial (popliteal) cyst, or synovial proliferation is readily identified on MRI. Subchondral cysts are easily detected on MRI because of the tomographic quality, multiplanar imaging capability, and the superb sensitivity to fluid- and fat-containing tissues. Cartilage pathology, both articular and meniscal, can be evaluated directly on MRI, and demonstration depends on the location of the abnormality and the pulse sequences used.

Magnetic resonance arthrography (MRA) performed with an intra-articular injection of dilute gadolinium contrast to improve cartilage evaluation has been investigated, but noncontrast MRI has been reported accurate for cartilage abnormalities. Patellofemoral cartilage loss has been reported to be closely associated with chronic knee pain symptoms.

Transient osteoporosis is characterized by self-limited pain and radiographically demonstrable osteopenia. The osteopenia typically develops within eight weeks after the onset of pain. Spontaneous osteonecrosis of the medial femoral condyle, most often found in middle-aged and elderly females, may have normal radiographs for months, followed by subchondral collapse, fragmentation of the articular cartilage, and progressive osteoarthritis. Bone marrow edema seen on MRI occurs in association with, or independent of, transient osteoporosis or osteonecrosis, and also in association with stress fractures; MRI is highly sensitive for detecting these abnormalities. In adult patients with conventional radiograph diagnosis of an osteochondral injury such as osteochondritis dissecans or osteonecrosis, an MRI examination may be indicated if an additional injury is suspected clinically or when it is necessary to determine the status of the articular cartilage over the area of abnormality. In the child or adolescent with radiographic evidence of osteochondritis dissecans, an MRI is indicated to determine the best method of treatment. Finally, MRI is not indicated to confirm a stress fracture that is evident on the radiographic study.

In patients with conventional evidence of inflammatory arthritis of the knee, the consensus of the panel is that a knee MRI is usually not indicated for preoperative differentiation of pannus from effusion or for evaluation of erosion. An aspiration for crystals may be indicated; however, the use of medical imaging (such as fluoroscopic guidance, ultrasound guidance, or arthrographic confirmation) is usually not necessary.

When intra-articular abnormality is suspected in a patient with claustrophobia, with a large body habitus, or who cannot, for some reason, tolerate an MRI examination, or when there is contraindication to an MRI, a computed tomography (CT) arthrogram may be used instead of the MRI to evaluate the cruciate ligaments, menisci, and articular cartilage.

**Summary**

The mandatory initial imaging examination for nontraumatic knee pain is AP and lateral radiography. In patients with anterior patellofemoral knee pain, an axial view should be included in the initial radiographic study. An MRI examination for nontraumatic knee pain is indicated when the pain is persistent and conventional radiographs are nondiagnostic or when additional information is necessary before instituting treatment or surgical intervention. An MRI is not indicated before a physical examination or routine conventional radiographs, or when there is diagnostic radiographic evidence of severe degenerative joint diseases, inflammatory arthritis, stress fracture, osteonecrosis, or reflex sympathetic dystrophy, for which additional imaging is not going to alter the treatment plan.

**Abbreviations**

- CT, computed tomography
- Med, medium
- Min, minimal
- MRI, magnetic resonance imaging
- NUC, nuclear medicine
- Tc, technetium
- US, ultrasound

<table>
<thead>
<tr>
<th>Relative Radiation Level</th>
<th>Effective Dose Estimated Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>0</td>
</tr>
<tr>
<td>Minimal</td>
<td>&lt;0.1 mSv</td>
</tr>
<tr>
<td>Low</td>
<td>0.1-1 mSv</td>
</tr>
<tr>
<td>Medium</td>
<td>1-10 mSv</td>
</tr>
<tr>
<td>High</td>
<td>10-100 mSv</td>
</tr>
</tbody>
</table>

**Clinical Algorithm(s)**

None provided

**Evidence Supporting the Recommendations**

**Type of Evidence Supporting the Recommendations**
The recommendations are based on analysis of the current literature and expert panel consensus.

Benefits/Harms of Implementing the Guideline Recommendations

**Potential Benefits**
Selection of appropriate radiologic imaging procedures to evaluate patients with nontraumatic knee pain

**Potential Harms**

*Relative Radiation Level (RRL)*
Potential adverse health effects associated with radiation exposure are an important factor to consider when selecting the appropriate imaging procedure. Because there is a wide range of radiation exposures associated with different diagnostic procedures, a relative radiation level indication has been included for each imaging examination. The RRLs are based on effective dose, which is a radiation dose quantity that is used to estimate population total radiation risk associated with an imaging procedure. Additional information regarding radiation dose assessment for imaging examinations can be found in the American College of Radiology (ACR) Appropriateness Criteria® Radiation Dose Assessment Introduction document (see "Availability of Companion Documents" field).

Qualifying Statements

An American College of Radiology (ACR) Committee on Appropriateness Criteria and its expert panels have developed criteria for determining appropriate imaging examinations for diagnosis and treatment of specified medical condition(s). These criteria are intended to guide radiologists, radiation oncologists, and referring physicians in making decisions regarding radiologic imaging and treatment. Generally, the complexity and severity of a patient’s clinical condition should dictate the selection of appropriate imaging procedures or treatments. Only those exams generally used for evaluation of the patient’s condition are ranked. Other imaging studies necessary to evaluate other co-existent diseases or other medical consequences of this condition are not considered in this document. The availability of equipment or personnel may influence the selection of appropriate imaging procedures or treatments. Imaging techniques classified as investigational by the U.S. Food and Drug Administration (FDA) have not been considered in developing these criteria; however, study of new equipment and applications should be encouraged. The ultimate decision regarding the appropriateness of any specific radiologic examination or treatment must be made by the referring physician and radiologist in light of all the circumstances presented in an individual examination.

Implementation of the Guideline

**Description of Implementation Strategy**
An implementation strategy was not provided.

Institute of Medicine (IOM) National Healthcare Quality Report Categories

**IOM Care Need**
- Getting Better
- Living with Illness

**IOM Domain**
- Effectiveness

Identifying Information and Availability

**Bibliographic Source(s)**

**Adaptation**
Not applicable: The guideline was not adapted from another source.

**Date Released**
1995 (revised 2008)

**Guideline Developer(s)**
American College of Radiology - Medical Specialty Society
Readers with questions regarding guideline content are directed to contact the guideline developer.
Maternal Fetal Medicine ACE
1. How would you rate your overall knowledge of Maternal Fetal Medicine?

<p>| | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>No knowledge</td>
<td>Excellent knowledge</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2. What are 3 goals you would like to achieve during the MFM ACE?
   a) 
   b) 
   c) 

3. What are 3 specific questions you would like to have answered during the MFM ACE?
   a) 
   b) 
   c)
MATERNAL FETAL MEDICINE

First Trimester US:

1. In an uncomplicated pregnancy (i.e. no pelvic pain, no bleeding, LMP known), after what gestational age is it optimal for when the first dating ultrasound should be performed?

2. At what gestational age can a fetal heart rate be detected on a transabdominal US?

3. At what gestational age can a fetal heart rate be detected on an endovaginal (EV) US?

First Trimester Screening (FTS):

1. What is a FTS?

2. In Alberta, who (i.e. at what maternal age) is eligible for a FTS?

3. Between what gestational ages should a FTS be performed?

4. What chromosome abnormalities are assessed with a FTS?

5. What information and tests are involved in a “comprehensive” FTS (circle all that apply)?

   Nuchal Translucency US   Bloodwork   Maternal Age   Nasal Bone on US

6. If bloodwork is performed as part of the FTS, is the “FTS bloodwork” equivalent to the “Maternal Serum Screen” (“MSS”)?

7. What 2 things are tested in “FTS bloodwork”?

8. For Trisomy 21 (Down Syndrome), what is detection rate (%) of Trisomy 21 (for a false positive rate of ~3%) based on:
   - Maternal age alone?
   - Maternal age + MSS at 15-18 weeks?
   - Maternal age + Nuchal Translucency US?
   - Maternal age + Nuchal Translucency US + FTS Bloodwork?
   - Maternal age + Nuchal Translucency US + Nasal Bone?
   - Maternal age + Nuchal Translucency US + Nasal Bone + FTS Bloodwork?

9. Can a “comprehensive” FTS be performed with a twin pregnancy?

Third Trimester US:

1. What are some common indications for a third trimester US?

Fetal Echocardiograms:

1. What are some indications for a screening fetal echocardiogram based on maternal and/or family history?

Recommended Reading:
## POST-ACE SELF-EVALUATION (MFM)

1. How would you rate your overall knowledge of Maternal Fetal Medicine *after the ACE*?

   1. No knowledge
   2.  
   3.  
   4.  
   5.  
   6.  
   7.  
   8.  
   9.  
   10. Excellent knowledge

2. Did you achieve the 3 goals you identified before the MFM ACE?

   ________________________________________________________________

3. What are 3 new things you learned from the MFM ACE?

   ________________________________________________________________
   ________________________________________________________________
   ________________________________________________________________

4. How will you change your practice after completing the MFM ACE?

   ________________________________________________________________
   ________________________________________________________________
   ________________________________________________________________
CanMEDS EVALUATION

(\textit{MFM})

Circle the non-Medical Expert CanMEDS Roles* addressed during the MFM ACE and give an example of how each Role was addressed.

*Adapted from 2005 The Royal College of Physicians and Surgeons of Canada (http://www.royalcollege.ca/shared/documents/canmeds/the_7_canmeds_roles_e.pdf)
OVERALL EVALUATION OF ACE
(MFM)

1. What is your overall impression of the MFM ACE?

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very Poor</td>
<td>Poor</td>
<td>Average</td>
<td>Very Good</td>
<td>Excellent</td>
</tr>
</tbody>
</table>

2. What were the 3 most helpful elements of the MFM ACE?

   a) __________________________________________

   b) __________________________________________

   c) __________________________________________

3. What are 3 elements of the MFM that were the least helpful?

   a) __________________________________________

   b) __________________________________________

   c) __________________________________________

4. Is there anything that you would like to be included in future MFM ACE’s?

   __________________________________________

   __________________________________________

   __________________________________________

5. Additional Comments?

   __________________________________________

   __________________________________________
Prenatal Screening for Fetal Aneuploidy in Singleton Pregnancies

This clinical practice guideline has been prepared by the Geneticians Committee of the Society of Obstetricians and Gynaecologists of Canada (SOGC) and the Prenatal Diagnosis Committee of the Canadian College of Medical Genetics (CCMG). It was approved by both the Executive and Council of the Society of Obstetricians and Gynaecologists of Canada and the Board of Directors of the Canadian College of Medical Genetics.

PRINCIPAL AUTHORS
David Chitayat, MD, Toronto ON
Sylvie Langlois, MD, Vancouver BC
R. Douglas Wilson, MD, Calgary AB

SOGC GENETICS COMMITTEE
R. Douglas Wilson, MD (Chair), Calgary AB
François Audibert, MD, Montreal QC
Claire Blight, RN, Halifax NS
Jo-Ann Brock, MD, Halifax NS
Lola Cartier, MSc, CCGC, Montreal QC
June Carroll, MD, Toronto ON
Valerie A. Désilets, MD, Montreal QC
Alain Gagnon, MD, Vancouver BC
Jo-Ann Johnson, MD, Calgary AB
Sylvie Langlois, MD, Vancouver BC
Lynn Murphy-Kauibeck, MD, Moncton NB
Nanette Okun, MD, Toronto ON
Melanie Pastuck, RN, Calgary AB
Vyta Senikas, MD, Ottawa ON

CCMG PRENATAL DIAGNOSIS COMMITTEE
Sylvie Langlois, MD (Chair), Vancouver BC
David Chitayat, MD, Toronto ON
Isabelle DeBie, MD, Montreal QC
Suzanne Demczuk, PhD, Saskatoon SK
Valerie A. Désilets, MD, Montreal QC
Michael T. Geraghty, MD, Ottawa ON
Janet Marcadier, MSc, Ottawa ON
Tanya N. Nelson, PhD, Vancouver BC
David Skidmore, MD, Halifax NS
Vicky Siu, MD, London ON

Disclosure statements have been received from all members of the committees.

Abstract
Objective: To develop a Canadian consensus document on maternal screening for fetal aneuploidy (e.g., Down syndrome and trisomy 18) in singleton pregnancies.

Options: Pregnancy screening for fetal aneuploidy started in the mid 1960s, using maternal age as the screening test. New developments in maternal serum and ultrasound screening have made it possible to offer all pregnant patients a non-invasive screening test to assess their risk of having a fetus with aneuploidy to determine whether invasive prenatal diagnostic testing is necessary. This document reviews the options available for non-invasive screening and makes recommendations for Canadian patients and health care workers.

Outcomes: To offer non-invasive screening for fetal aneuploidy (trisomy 13, 18, 21) to all pregnant women. Invasive prenatal diagnosis would be offered to women who screen above a set risk cut-off level on non-invasive screening or to pregnant women whose personal, obstetrical, or family history places them at increased risk. Currently available non-invasive screening options include maternal age combined with one of the following: (1) first trimester screening (nuchal translucency, maternal age, and maternal serum biochemical markers), (2) second trimester serum screening (maternal age and maternal serum biochemical

Key Words: Aneuploidy, Down syndrome, trisomy, prenatal screening, genetic health risk, genetic health surveillance, prenatal diagnosis

This document reflects emerging clinical and scientific advances on the date issued, and is subject to change. The information should not be construed as dictating an exclusive course of treatment or procedure to be followed. Local institutions can dictate amendments to these opinions. They should be well documented if modified at the local level. None of these contents may be reproduced in any form without prior written permission of the SOGC.
markers), or (3) 2-step integrated screening, which includes first and second trimester serum screening with or without nuchal translucency (integrated prenatal screen, serum integrated prenatal screening, contingent, and sequential). These options are reviewed, and recommendations are made.

Evidence: Studies published between 1982 and 2009 were retrieved through searches of PubMed or Medline and CINAHL and the Cochrane Library, using appropriate controlled vocabulary and key words (aneuploidy, Down syndrome, trisomy, prenatal screening, genetic health risk, genetic health surveillance, prenatal diagnosis).

Results were restricted to systematic reviews, randomized controlled trials, and relevant observational studies. There were no language restrictions. Searches were updated on a regular basis and incorporated in the guideline to August 2010. Grey (unpublished) literature was identified through searching the websites of health technology assessment and health technology assessment-related agencies, clinical practice guideline collections, clinical trial registries, and national and international medical specialty societies.

The previous Society of Obstetricians and Gynaecologists of Canada guidelines regarding prenatal screening were also reviewed in developing this clinical practice guideline.

Values: The quality of evidence was rated using the criteria described in the Report of the Canadian Task Force on Preventive Health Care.

Benefits, harms, and costs: This guideline is intended to reduce the number of prenatal invasive procedures done when maternal age is the only indication. This will have the benefit of reducing the numbers of normal pregnancies lost because of complications of invasive procedures. Any screening test has an inherent false-positive rate, which may result in undue anxiety. It is not possible at this time to undertake a detailed cost-benefit analysis of the implementation of this guideline, since this would require health surveillance and research and health resources not presently available; however, these factors need to be evaluated in a prospective approach by provincial and territorial initiatives.

Recommendations
1. All pregnant women in Canada, regardless of age, should be offered, through an informed counselling process, the option of a prenatal screening test for the most common clinically significant fetal aneuploidies in addition to a second trimester ultrasound for dating, assessment of fetal anatomy, and detection of multiples. (I-A)

2. Counselling must be non-directive and must respect a woman’s right to accept or decline any or all of the testing or options offered at any point in the process. (III-A)

3. Maternal age alone is a poor minimum standard for prenatal screening for aneuploidy, and it should not be used a basis for recommending invasive testing when non-invasive prenatal screening for aneuploidy is available. (II-2A)

4. Invasive prenatal diagnosis for cytogenetic analysis should not be performed without multiple marker screening results except for women who are at increased risk of fetal aneuploidy (a) because of ultrasound findings, (b) because the pregnancy was conceived by in vitro fertilization with intracytoplasmic sperm injection, or (c) because the woman or her partner has a history of a previous child or fetus with a chromosomal abnormality or is a carrier of a chromosome rearrangement that increases the risk of having a fetus with a chromosomal abnormality. (II-2E)

5. At minimum, any prenatal screen offered to Canadian women who present for care in the first trimester should have a detection rate of 75% with no more than a 3% false-positive rate. The performance of the screen should be substantiated by annual audit. (III-B)

6. The minimum standard for women presenting in the second trimester should be a screen that has a detection rate of 75% with no more than a 5% false-positive rate. The performance of the screen should be substantiated by annual audit. (III-B)

7. First trimester nuchal translucency should be interpreted for risk assessment only when measured by sonographers or sonologists trained and accredited for this service and when there is ongoing quality assurance (II-2A), and it should not be offered as a screen without biochemical markers in singleton pregnancies. (I-E)

8. Evaluation of the fetal nasal bone in the first trimester should not be incorporated as a screen unless it is performed by sonographers or sonologists trained and accredited for this service and there is ongoing quality assurance. (II-2E)

9. For women who undertake first trimester screening, second trimester serum alpha fetoprotein screening and/or ultrasound examination is recommended to screen for open neural tube defects. (II-1A)

10. Timely referral and access is critical for women and should be facilitated to ensure women are able to undergo the type of screening test they have chosen as first trimester screening. The first steps of integrated screening (with or without nuchal translucency), contingent, or sequential screening are performed in an early and relatively narrow time window. (II-1A)

11. Ultrasound dating should be performed if menstrual or conception dating is unreliable. For any abnormal serum screen calculated on the basis of menstrual dating, an ultrasound should be done to confirm gestational age. (II-1A)

12. The presence or absence of soft markers or anomalies in the 18- to 20-week ultrasound can be used to modify the a priori risk of aneuploidy established by age or prior screening. (II-2B)

13. Information such as gestational dating, maternal weight, ethnicity, insulin-dependent diabetes mellitus, and use of assisted reproduction technologies should be provided to the laboratory to improve accuracy of testing. (II-2A)

14. Health care providers should be aware of the screening modalities available in their province or territory. (III-B)

15. A reliable system needs to be in place ensuring timely reporting of results. (III-C)

ABBREVIATIONS

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFP</td>
<td>alpha fetoprotein</td>
</tr>
<tr>
<td>CVS</td>
<td>chorionic villus sampling</td>
</tr>
<tr>
<td>DR</td>
<td>detection rate</td>
</tr>
<tr>
<td>FPR</td>
<td>false-positive rate</td>
</tr>
<tr>
<td>FTS</td>
<td>first trimester screening</td>
</tr>
<tr>
<td>hCG</td>
<td>human chorionic gonadotropin</td>
</tr>
<tr>
<td>IPS</td>
<td>integrated prenatal screening</td>
</tr>
<tr>
<td>MMS</td>
<td>multiple marker screening</td>
</tr>
<tr>
<td>MoM</td>
<td>multiples of the median</td>
</tr>
<tr>
<td>MSAFP</td>
<td>maternal serum alpha fetoprotein</td>
</tr>
<tr>
<td>NT</td>
<td>nuchal translucency</td>
</tr>
<tr>
<td>ONTD</td>
<td>open neural tube defect</td>
</tr>
<tr>
<td>PAPP-A</td>
<td>pregnancy-associated plasma protein-A</td>
</tr>
<tr>
<td>PR</td>
<td>positive rate</td>
</tr>
<tr>
<td>SLOS</td>
<td>Smith-Lemli-Opitz Syndrome</td>
</tr>
<tr>
<td>uE3</td>
<td>unconjugated estriol</td>
</tr>
</tbody>
</table>
16. Screening programs should be implemented with resources that support audited screening and diagnostic laboratory services, ultrasound, genetic counselling services, patient and health care provider education, and high quality diagnostic testing, as well as resources for administration, annual clinical audit, and data management. In addition, there must be the flexibility and funding to adjust the program to new technology and protocols. (II-3B)

J Obstet Gynaecol Can 2011;33(7):736–750

INTRODUCTION

Screening for chromosomal anomalies and open neural tube defects is part of prenatal care offered to all Canadian women. Since the methods of screening for ONTDs have not changed since the mid-1970s, they are not discussed here. Screening for ONTDs in Canada requires second trimester serum alpha fetoprotein (16 to 20 completed weeks) and/or ultrasound examination done at 18 to 22 weeks of gestation.

Screening for fetal chromosomal anomalies, including Down syndrome, began with amniocentesis in the mid-1960s. At that time, the criterion for screening was maternal age. In Canada, screening was offered only to women ≥ 35 years at the expected date of delivery. This was determined to be the point at which the risk of a pregnancy loss was less than the chance of identifying a pregnancy with a significant chromosomal abnormality. This clinical practice guideline reviews the evolution of screening for fetal aneuploidy from screening using maternal age alone to the many options currently available and makes recommendations regarding the minimum standard of prenatal screening that should be available to all Canadian women. The level of evidence and quality of recommendations are described using the criteria and classifications of the Canadian Task Force on Preventive Health Care (Table 1).

WHAT IS SCREENING?

Screening is the process of surveying a population, using a specific marker or markers and defined screening cut-off levels, to identify the individuals in the population at higher risk for a particular disorder. Screening is applicable to a population; diagnosis is applied at the individual patient level.

Screening for a disorder should be undertaken only when the disorder is considered to be serious enough to warrant intervention. The marker or markers used in screening must be sufficiently sensitive to identify a significant proportion of affected persons with minimal misidentification of unaffected persons. There must also be both a highly accurate diagnostic test to determine whether the person with a screen positive result truly has the disorder and an intervention available to all persons who are identified as being affected. The screen, including follow-up testing and intervention, must be affordable. Finally the screen must be acceptable to the population being screened.

The screening procedure should not be merely a test but must be a comprehensive program. The program must include the provision of understandable information for both patients and health care providers to ensure informed decision-making, timely access to the screening test, a system of notification of results and referral to follow-up testing, and access to an intervention. The screening process must allow patients to decline intervention at each step throughout the process. A screening program must include regular clinical audit to evaluate local performance and should have the flexibility to incorporate new technology.

The Appendix provides a glossary of terms commonly used in screening

IMPORTANT CONCEPTS UNDERLYING PRENATAL GENETIC SCREENING

Multiple marker screening uses a combination of maternal age and 2 or more biochemical tests, with or without an ultrasound examination, to produce a single result for risk of Down syndrome, trisomy 18, and ONTDs, which is used to offer options for clinical management. A screen is positive when the risk of one or more of the screened disorders falls above a designated risk cut-off. Counselling and further testing options are offered when a screen is positive. In the discussion of prenatal screening, the terms false-positive rate or positive rate, and detection rate are used (Appendix). As screening performance improves, the FPR decreases and/or the DR increases. A risk cut-off might be chosen based upon the desired DR, FPR, or a combination of both. A risk cut-off is expressed as the risk or likelihood of the condition being present in the fetus at term or at mid-trimester. The risk for the latter will be higher, because 23% of fetuses with Down syndrome are miscarried between mid-trimester and term (risk cut-off of 1:350 at term would be similar to 1:280 at mid-trimester).

The other commonly used term in multiple marker screening is multiples of the median. Each marker result, including both biochemistry and nuchal translucency measurements, can be expressed in MoM. The absolute value of the assayed marker (serum or nuchal translucency) is divided by the gestation-specific median value of the serum marker in the measuring laboratory or by using standard or sonographer-specific curves for nuchal translucency. This allows direct comparison of results between programs.
Table 1. Key to evidence statements and grading of recommendations, using the ranking of the Canadian Task Force on Preventive Health Care

<table>
<thead>
<tr>
<th>Quality of evidence assessment*</th>
<th>Classification of recommendations†</th>
</tr>
</thead>
<tbody>
<tr>
<td>I: Evidence obtained from at least one properly randomized controlled trial</td>
<td>A. There is good evidence to recommend the clinical preventive action</td>
</tr>
<tr>
<td>II-1: Evidence from well-designed controlled trials without randomization</td>
<td>B. There is fair evidence to recommend the clinical preventive action</td>
</tr>
<tr>
<td>II-2: Evidence from well-designed cohort (prospective or retrospective) or case–control studies, preferably from more than one centre or research group</td>
<td>C. The existing evidence is conflicting and does not allow to make a recommendation for or against use of the clinical preventive action; however, other factors may influence decision-making</td>
</tr>
<tr>
<td>II-3: Evidence obtained from comparisons between times or places with or without the intervention. Dramatic results in uncontrolled experiments (such as the results of treatment with penicillin in the 1940s) could also be included in this category</td>
<td>D. There is fair evidence to recommend against the clinical preventive action</td>
</tr>
<tr>
<td>III: Opinions of respected authorities, based on clinical experience, descriptive studies, or reports of expert committees</td>
<td>E. There is good evidence to recommend against the clinical preventive action</td>
</tr>
<tr>
<td></td>
<td>F. There is insufficient evidence (in quantity or quality) to make a recommendation; however, other factors may influence decision-making</td>
</tr>
</tbody>
</table>

* The quality of evidence reported in these guidelines has been adapted from The Evaluation of Evidence criteria described in the Canadian Task Force on Preventive Health Care.
† Recommendations included in these guidelines have been adapted from the Classification of Recommendations criteria described in the Canadian Task Force on Preventive Health Care.

SCREENING FOR CHROMOSOMAL ANOMALIES

Traditionally, in Canada, the option of invasive testing was recommended when a woman's risk of having a pregnancy with a chromosome anomaly was higher than the risks associated with the common invasive procedures (amniocentesis or CVS). New developments in maternal serum and ultrasound screening have improved the ability to identify pregnancies at increased risk for Down syndrome, trisomy 18, and other chromosomal abnormalities. This allows use of these screening tests to identify pregnancies at high enough risk to warrant invasive diagnostic testing, which has a risk of pregnancy loss.

The most common chromosome conditions associated with advanced maternal age involve the presence of an additional chromosome (21, 18, 13, or X). Down syndrome, trisomy 18, and trisomy 13 are associated with congenital anomalies and intellectual disability. With ultrasound and maternal serum screening, pregnancies affected by these conditions can now be recognized with a significant degree of reliability. The practice of using solely maternal age at the estimated date of delivery to identify at-risk pregnancies should be abandoned. The maternal age-related risk should be modified by additional non-invasive markers, which consist of maternal serum markers and ultrasound assessment. The approach to screening and diagnosis may vary between provinces. It is the responsibility of each health care provider to be aware of what is available to his or her patients so that appropriate counselling is provided.

INVASIVE PREGNATAL DIAGNOSIS TO BE LIMITED TO WOMEN AT INCREASED RISK OF FETAL ANEUPLOIDY

The probability of conceiving a fetus with a trisomy increases with maternal age. Prenatal screening for chromosome anomalies starts with a discussion of the maternal age-related risk of having a baby with a chromosomal abnormality. However, maternal age screening is inferior to newer screening approaches that use multiple biochemical markers with or without a first trimester ultrasound assessment of nuchal translucency. These strategies provide a greatly reduced FPR with a substantially improved DR when applied across all age groups, and they provide evidence that the practice of offering invasive prenatal diagnosis for age alone as an indication should be abandoned.³ Women ≥ 40 years should not be offered an amniocentesis without prior screening, because with a negative screening result, their risk of a clinically significant chromosomal abnormality remains < 1/200. Invasive prenatal diagnosis for cytogenetic analysis should be offered only to women who are considered to be at increased risk of fetal aneuploidy on the basis of a non-invasive screen result above the risk cut-off, because of ultrasound findings, because the pregnancy was conceived by IVF with intracytoplasmic sperm injection,⁴ or because the woman or her partner has a history of a previous child or fetus with a chromosomal abnormality or is a carrier of a chromosome rearrangement that increases the risk of having a fetus with a chromosomal abnormality.
In these scenarios, the risk of a chromosomal abnormality, including chromosomal anomalies not detected by screening, is high enough to offer invasive testing without prior screening.

**Recommendations**

1. All pregnant women in Canada, regardless of age, should be offered, through an informed counselling process, the option of a prenatal screening test for the most common clinically significant fetal aneuploidies in addition to a second trimester ultrasound for dating, assessment of fetal anatomy, and detection of multiples (I-A)

2. Counselling must be non-directive and must respect a woman’s right to accept or decline any or all of the testing or options offered at any point in the process. (II-A)

3. Maternal age alone is a poor minimum standard for prenatal screening for aneuploidy, and it should not be used as a basis for recommending invasive testing when non-invasive prenatal screening for aneuploidy is available. (II-2A)

4. Invasive prenatal diagnosis for cytogenetic analysis should not be performed without multiple marker screening results except for women who are at increased risk of fetal aneuploidy (a) because of ultrasound findings, (b) because the pregnancy was conceived by in vitro fertilization with intracytoplasmic sperm injection, or (c) because the woman or her partner has a history of a previous child or fetus with a chromosomal abnormality or is a carrier of a chromosome rearrangement that increases the risk of having a fetus with a chromosomal abnormality. (II-2E)

**CHOOSING A SCREEN**

The most appropriate screening test for Down syndrome would have the lowest FPR and the highest DR. Cost and logistics should also be considered. Generally, the costs associated with screening are measured by the cost per Down syndrome pregnancy diagnosed. This has been estimated using different screening options in several studies.5-9 One of the difficulties with cost analyses is that the expenses associated with the ultrasound and serum sample analyses vary greatly from one jurisdiction to another. In addition, cost has not been estimated for many screening options, including the second trimester ultrasound. Consequently, a comprehensive cost comparison remains to be undertaken.

Given geographic limitations and resource differences, it is unlikely that a single screening protocol can be endorsed or practically applied for all women across Canada; however, screening options should meet acceptable performance characteristics. Considering the tests currently available and the risk and benefit ratio, it is believed that at a minimum, screening programs should provide a screen that offers women who present in the first trimester a DR for Down syndrome of 75% with no more than a 3% FPR.10,11 For women presenting in the second trimester, the screen offered should have a minimum DR of 75% with no more than a 5% FPR. Table 2 provides details of currently available screening options and their screening performance. Table 3 details timing of results for options that meet the minimum standard. These include first trimester screening, quad screening in second trimester, 2-step integrated first and second trimester prenatal serum screening with or without nuchal transluency (IPS and serum IPS).12 IPS can be offered as full integrated screening for all women or as contingent or sequential screening. Access to follow-up services should also be ensured. Finally, prenatal screening programs must balance minimizing the FPR (which minimizes the number of invasive procedures needed and thus the number of normal pregnancies lost to chorionic villus sampling or amniocentesis) against the desire to detect as many cases as possible as early in gestation as possible. Some studies suggest women prefer a lower positive rate,13-15 while others suggest that women want early diagnosis.16,17 Individual programs should determine what is appropriate for their jurisdiction.

5. At minimum, any prenatal screen offered to Canadian women who present for care in the first trimester should have a detection rate of 75% with no more than a 3% false-positive rate. The performance of the screen should be substantiated by annual audit. (III-B)

6. The minimum standard for women presenting in the second trimester should be a screen that has a detection rate of 75% with no more than a 5% false-positive rate. The performance of the screen should be substantiated by annual audit. (III-B)

**REVIEW OF SCREENING OPTIONS**

First Trimester Screening: Nuchal Translucency Combined With Biochemical Markers

Nuchal translucency refers to the subcutaneous layer of fluid behind the fetal neck and lower cranium, which can be visualized on ultrasound. In 1992, Nicolaides et al.18
Table 2. Current available screening options and their screening performance

<table>
<thead>
<tr>
<th>Screening option</th>
<th>Markers</th>
<th>Trimester</th>
<th>Term risk cut-off</th>
<th>DR, %</th>
<th>FPR, %</th>
<th>OAPR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Options that meet the minimum standard:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FTS&lt;sup&gt;1,10&lt;/sup&gt;</td>
<td>NT, free β-hCG, PAPP-A, MA</td>
<td>1st</td>
<td>1 in 325</td>
<td>83</td>
<td>5.0</td>
<td>1:27</td>
</tr>
<tr>
<td>Quad screening&lt;sup&gt;11&lt;/sup&gt;</td>
<td>AFP, uE3, free β-hCG, inhibin A, MA</td>
<td>2nd</td>
<td>1 in 385</td>
<td>77</td>
<td>5.2</td>
<td>1:50</td>
</tr>
<tr>
<td>IPS&lt;sup&gt;7,10&lt;/sup&gt;</td>
<td>NT, PAPP-A, AFP, uE3, free β-hCG/total hCG, inhibin A, MA</td>
<td>1st &amp; 2nd</td>
<td>1 in 200</td>
<td>87</td>
<td>1.9</td>
<td>1:10</td>
</tr>
<tr>
<td>IPS without inhibin A&lt;sup&gt;2&lt;/sup&gt;</td>
<td>NT, PAPP-A, AFP, uE3, total hCG, MA</td>
<td>1st &amp; 2nd</td>
<td>1 in 200</td>
<td>88</td>
<td>3.0</td>
<td>1:20</td>
</tr>
<tr>
<td>Serum IPS&lt;sup&gt;7,10&lt;/sup&gt;</td>
<td>PAPP-A, AFP, uE3, free β-hCG/total hCG, inhibin A</td>
<td>1st &amp; 2nd</td>
<td>1 in 200</td>
<td>85</td>
<td>4.4</td>
<td>1:26</td>
</tr>
<tr>
<td>Options that do not meet the minimum standard:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maternal age&lt;sup&gt;12&lt;/sup&gt;</td>
<td>MA</td>
<td>1st &amp; 2nd</td>
<td>1 in 385</td>
<td>44</td>
<td>16</td>
<td>1:218</td>
</tr>
<tr>
<td>Triple screening&lt;sup&gt;12&lt;/sup&gt;</td>
<td>AFP, uE3, total hCG, MA</td>
<td>2nd</td>
<td>1 in 385</td>
<td>71</td>
<td>7.2</td>
<td>1:59</td>
</tr>
</tbody>
</table>

MA: Maternal age. OAPR: Odds of being affected given a positive result.

* Some centres in Canada may offer variation on IPS (sequential screening or contingent screening) with cut-offs set that achieve at least the minimum standard.

Table 3. Available screening options that meet minimum standard

<table>
<thead>
<tr>
<th>Screening methods that meet guideline minimal standard of 75% DR with 3% FPR</th>
<th>Timing of results</th>
<th>Is 2nd trimester ultrasound still recommended?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st trimester screen</td>
<td>1st trimester</td>
<td>Yes</td>
</tr>
<tr>
<td>2nd trimester quad screen</td>
<td>2nd trimester</td>
<td>Yes</td>
</tr>
<tr>
<td>Two-step screens</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contingent</td>
<td>For most patients, result available in 1st trimester; small proportion of patients require second trimester testing</td>
<td>Yes</td>
</tr>
<tr>
<td>Integrated</td>
<td>Single result in 2nd trimester</td>
<td>Yes</td>
</tr>
<tr>
<td>Serum integrated</td>
<td>Single result in 2nd trimester</td>
<td>Yes</td>
</tr>
<tr>
<td>Sequential</td>
<td>Results in 1st and 2nd trimester for the same patient</td>
<td>Yes</td>
</tr>
</tbody>
</table>

described an association between an increased size of the nuchal translucency on the 11- to 14-week fetal ultrasound scan and an increased risk of fetal Down syndrome. Several large studies have shown that NT has a DR for Down syndrome ranging from 69% to 75%, with an FPR of 5% to 8.1%.<sup>5,10,19</sup> This marker is associated with other numeric chromosome abnormalities, other fetal anomalies such as cardiac defects or diaphragmatic hernia, and a number of single gene disorders, particularly those associated with decreased fetal movement. An NT above the 99th percentile has a sensitivity of 31% and specificity of 98.7% for major congenital heart defects when the fetal karyotype is normal. One in 33 fetuses with an NT above the 95th percentile (above 2.2 to 2.8 mm depending on gestational age) and 1 in 16 with an NT above the 99th percentile (≥ 3.5 mm regardless of gestational age) have a major cardiac defect detected.<sup>20</sup> Finding an increased NT at 11 to 14 weeks’ gestation with a normal fetal karyotype warrants offering a detailed ultrasound examination at 18 to 20 weeks, with an assessment of the fetal heart, including a 4-chamber view and view of the outflow tracts as a minimum<sup>21</sup> or a fetal echocardiogram if available.

Two first trimester maternal serum biochemical markers emerged at the same time as NT was being investigated. These are PAPP-A and hCG (free beta or total). PAPP-A is lower in Down syndrome pregnancies and hCG is higher.<sup>22,23</sup> In a study by Wald and Hackshaw that used a combination of the maternal age-related risk, maternal serum PAPP-A and free β-hCG, the DR of Down syndrome was 61%, with a 5% false-positive rate.<sup>24</sup> The first trimester biochemical markers alone were not as efficacious as second trimester screening; however, a combination of the 2 first trimester biochemical markers with NT, demonstrated a significant improvement over second trimester triple and quadruple screening. FTS using maternal age, NT plus PAPP-A, and free β-hCG will detect 83% of cases of Down syndrome, with a 5% FPR,
using a term risk cut-off for Down syndrome of 1:300 (or will detect 78% of cases of Down syndrome with a 3% FPR) and thus fulfills the guideline recommendation. FTS also has the ability to screen for trisomies 13 and 18.

Limitations on using FTS include the availability and reproducibility of NT as well as the availability of chorionic villus sampling as a diagnostic testing option for those with a screen positive result. Guidelines for measuring NT to maximize reproducibility and accuracy have been developed by the Fetal Medicine Foundation in the United Kingdom. The Royal College of Obstetricians and Gynaecologists (United Kingdom) study group on first trimester assessment of Down syndrome recommended that NT measurement should be implemented only in centres with appropriately trained sonographers using high-quality equipment, and that the results should be subject to regular audit by an external agency. The use of NT in a clinical setting requires a program of quality control and maintenance of skills through an ongoing audit of NT measurements to achieve standardization and maintain the quality essential to obtain the desired DR and FPR of the screening tests.

Finally, if local ultrasound services are unable to provide a comprehensive screen for neural tube defects at 18 to 20 weeks’ gestation, patients undergoing first trimester screening for aneuploidy should be offered MSAFP in the second trimester to screen for open neural tube defects. Ultrasound screening for delayed ossification of the fetal nasal bone can be done in the first or second trimester. The first trimester ultrasound, which determines the presence or absence of the nasal bone between 11 and 14 weeks of gestation, may be more likely to be incorporated into other screening modalities. First trimester assessment of the fetal nasal bone was described by Ciocero et al. and detected 77% of Down syndrome cases. Subsequent work has shown a DR of 68.8% and that the FPR depends upon maternal ethnicity (9% in Afro-Caribbeans, 5% in Asians, and 2.2% in Caucasians). The FPR also varied with crown-rump length (increasing with decreasing crown-rump length) and NT (increasing with increasing NT). The difficulty in performing first trimester nasal bone sonography consistently in the general population might limit the usefulness of this screening technique. A study of intra- and interoperator variability in fetal nasal bone assessment during the first trimester showed that the assessment was only fairly reproducible. Guidelines for the ultrasound assessment of the nasal bone have been developed by the Fetal Medicine Foundation. As with the use of NT, ultrasound assessment of the nasal bone in a clinical setting requires a program of training, quality control, and maintenance of skills through annual audit of nasal bone images.

### Recommendations

7. First trimester nuchal translucency should be interpreted for risk assessment only when measured by sonographers or sonologists trained and accredited for this service and when there is ongoing quality assurance (II-2A), and it should not be offered as a screen without biochemical markers in singleton pregnancies. (I-E)

8. Evaluation of the fetal nasal bone in the first trimester should not be incorporated as a screen unless it is performed by sonographers or sonologists trained and accredited for this service and there is ongoing quality assurance. (II-2E)

9. For women who undertake first trimester screening, second trimester serum alpha fetoprotein screening and/or ultrasound examination is recommended to screen for open neural tube defects. (II-1A)

### Second Trimester Screening

In the 1970s, alpha fetoprotein was identified as a second trimester marker for open neural tube defects. MSAFP continues to be used as part of multiple marker screening for this purpose, but is also effective as a screen for other open fetal defects such as gastroschisis and omphalocele.

In 1983, low MSAFP was noted in a patient who had a baby with trisomy 18. Further investigation showed this marker was low in Down syndrome as well, and for a few years, MSAFP combined with maternal age was used as a marker for Down syndrome. In 1988, Wald et al. demonstrated that the combination of maternal age and MSAFP with two other maternal serum markers, unconjugated estriol (MSue3) and human chorionic gonadotrophin (MSHCG), measured between 15 and 20 weeks’ gestation, would detect 65% of fetuses with Down syndrome using a 5% FPR. These predictions were confirmed in several subsequent studies. Triple marker screening has been available in Canada since 1991. Using a term risk cut-off of 1:385, the triple marker screening detects 72% of fetuses with Down syndrome with a 7% FPR. The triple marker screening also screens for ONTDs, other open fetal defects (e.g., gastroschisis, omphalocele), placental dysfunction, Smith-Lemli-Opitz syndrome, and trisomy 18. The triple screen does not fulfill the guideline recommendation.

Inhibin A is a fourth marker that can be added in the second trimester, resulting in the quad screen. Inhibin A will increase the DR of Down syndrome by 10%. With a risk cut-off of 1:230 at term, the DR is 75% to 80%, and the FPR is lowered to 3% to 5%, thus meeting the minimal standard set by this guideline.
Combined First and Second Trimester Options
Integrated prenatal screening

In an effort to further improve performance, the first and second trimester screening tests have been combined into a process called integrated prenatal screening. Wald et al.37 predicted that integrating first and second trimester screening would result in an 83% DR for Down syndrome, with a 2.1% FPR at a term risk cut-off of 1:200. IPS was based on the use of PAPP-A and NT in the first trimester and the quad screen in the second trimester, with results released when all the testing was completed.37 This approach has been controversial, with some authors suggesting women had the right to know their results early and that it was unethical to withhold the first trimester results.38 However, when IPS utilizes a quad screen in the second trimester, studies have shown a detection rate of 85% to 87% with an FPR of 0.8% to 1.5%.5.10 When Inhibin A is excluded from the IPS, the FPR increases to ~2.5% when the first trimester markers are performed at 12 weeks. Full integrated screening meets the guideline minimal standard. The benefit of IPS over FTS is the achievement of a lower FPR and reduction of the number of invasive diagnostic procedures needed.

The optimal time for the PAPP-A measurement is 9 to 10 weeks’ gestation with the performance of PAPP-A decreasing between 10 and 13 weeks. The proportion of pregnancies in which a satisfactory NT measurement can be obtained is the highest at 11 to 13 weeks’ gestation. First trimester measurements are usually carried out between 11 and 14 weeks’ gestation as a compromise to make the timing favourable for NT and PAPP-A.5 IPS also screens for open fetal neural tube defects and trisomy 18.

Serum Integrated prenatal screening

When NT is not available, IPS still can be offered, using PAPP-A in the first trimester and triple or quad screening in the second trimester. This approach has an 83% DR for a 4% FPR.5 Alternatively, PAPP-A and free β-hCG can be offered in the first trimester, followed by AFP and uE3 in the second with virtually the same performance. The FPR is 4.2% if PAPP-A is measured at 10 completed weeks, and the FPR is doubled (8.5%) if it is measured at 13 completed weeks.5 In the FASTER trial, serum IPS showed a 4.4% FPR for an 85% DR.10 Serum IPS is a practical option for areas of Canada where there is limited or no access to NT screening.

Given that timing is critical for serum analysis, accurate dating of the pregnancy is very important. Ultrasound dating should be performed if menstrual or conception dating is unreliable. For any abnormal serum screen (serum IPS, quad) calculated using menstrual dating, an ultrasound should be done to confirm gestational age.

Recommendations

10. Timely referral and access is critical for women and should be facilitated to ensure women are able to undergo the type of screening test they have chosen as first trimester screening. The first steps of integrated screening (with or without nuchal translucency), contingent, or sequential screening are performed in an early and relatively narrow time window. (II-1A)

11. Ultrasound dating should be performed if menstrual or conception dating is unreliable. For any abnormal serum screen calculated on the basis of menstrual dating, an ultrasound should be done to confirm gestational age. (II-1A)

Contingent screening

The concept of contingent screening has been suggested by Wright et al.46 as an alternative to IPS. In contingent screening, the majority of women receive their result after FTS. Women at high risk (risk >1/50) are offered invasive testing, and women at low risk (risk <1/1500) require no further testing. A proportion of women with a risk between two cut-offs (1/50 and 1/1500) will go on to have second trimester screening and will receive a combined result. Benn et al.47 reported the expected screening performance of the contingent strategy by modelling on different risk cut-offs and maternal age distributions of the United Kingdom and the United States. Performance of contingent screening was comparable with IPS if total hCG and/or free β-hCG was measured in both trimesters.48 It is possible to select risk cut-offs that achieve performances similar to IPS, thus meeting the guideline recommendation, while achieving detection of a significant proportion of abnormal pregnancies by the end of the first trimester.41,42 A study of computer simulations to compare integrated, sequential, and contingent screening strategies with various cut-offs leading to 19 potential screening algorithms showed that the contingent screening strategy had the best cost-effectiveness ratio, with fewer procedure-related euploid miscarriages and unnecessary terminations.43 However, in contingent screening, a proportion of women are identified as having an intermediate risk and asked to have the second trimester serum to modify their risk. Using a prospective first trimester cohort of 18 901 pregnancies and a contingent protocol, Coccioletone et al. reported that 15.8% of cases had first trimester combined risk odds between 1 per 51 and 1 per 1500, thus requiring second trimester serum marker analysis.44 The women in this intermediate risk group are likely to experience raised anxiety, and a proportion of them might wish to have invasive testing immediately, thus increasing the FPR.42,45
Sequential screening
Sequential screening selects women for second trimester testing on the basis of their first trimester screening results. Women found to be at high risk on the basis of the FTS (e.g., risk \( \geq 1 \) in 50) are offered invasive testing. Those with a risk lower than the cut-off are offered additional serum screening in the second trimester. The removal of screen positive affected cases in the first trimester decreases the prevalence of Down syndrome in the second trimester and consequently lowers the PPV of second trimester screening.\(^{46,47}\) As a result the overall cut-off is adjusted to take this into consideration. With appropriate cut-offs, sequential screening has been shown to perform equivalently to full integrated and contingent screening and meets the guideline recommendation.\(^{54}\)

Sequential screening that does not incorporate the results of the first trimester testing into the second trimester risk analysis is associated with a significant increased FPR.\(^{46,47}\) Given this high FPR, sequential screening should not be offered unless the second trimester risk incorporates the first trimester results.

POTENTIAL OF SCREENING OPTIONS TO DETECT CHROMOSOMAL ANOMALIES OTHER THAN DOWN SYNDROME AND OTHER GENETIC CONDITIONS

In pregnancies with trisomy 18, first trimester PAPP-A is decreased, NT is enlarged, and second trimester serum levels of AFP, uE3, hCG, and inhibit A are significantly reduced.\(^{48-51}\) Many centres are now routinely screening for trisomy 18, using protocols designed for this anomaly. With second trimester triple marker screening, at a term risk cut-off of \( \geq 1:100 \), 60% of trisomy 18 pregnancies can be detected for a FPR of 0.2%.\(^{52}\) With serum IPS, using the same cut-off, the DR is 90% for a FPR of 0.1%.\(^{53}\) A protocol for the detection of trisomies 13 and 18 has been developed for FTS.\(^{54}\)

Studies show a large proportion of fetuses with triploidy can be detected with the current MSS or FTS protocols.\(^{55,56}\) Second trimester uE3 is decreased and inhibit A is elevated in pregnancies with trisomy 13.\(^{37,38}\) Turner syndrome is associated with a lower uE3. Higher hCG and inhibit A levels also are seen in cases where there is fetal hydrops.\(^{59-61}\) Increased NT and a lower PAPP-A have been reported in pregnancies with triploidy of paternal origin, trisomy 13, Turner syndrome, and other sex chromosome aneuploidies.\(^{62-64}\) Trisomy 13 and 18, Turner syndrome, and triploidy also are associated with anomalies and markers that allow the majority to be detected during the 18- to 20-week ultrasound.

Smith-Lemli-Opitz Syndrome is an autosomal recessive disorder associated with intellectual disability and multiple congenital anomalies. The minimum incidence is estimated to be 1 in 60 000.\(^{65}\) SLOS is due to an abnormality in cholesterol synthesis resulting in a low cholesterol concentration and accumulation of its precursors in blood and tissue.\(^{66}\) SLOS can be diagnosed prenatally by the presence of abnormally elevated amniotic fluid 7-dehydrocholesterol concentration.\(^{67}\) In pregnancies with SLOS, maternal serum uE3, AFP, and hCG are reduced.\(^{68}\) A screening protocol has been developed for this syndrome that provides a DR of 62% for a FPR of 0.33%.\(^{69}\) However, the screen is not specific for SLOS since it detects a number of rare disorders of cholesterol and estriol biosynthesis, such as congenital adrenal hypoplasia and Zellweger syndrome, as well as a relatively common and mild disorder, X-linked steroid sulfatase deficiency (X-linked ichthyosis).\(^{70}\)

THE USE OF ULTRASOUND IN SCREENING FOR CHROMOSOMAL ANOMALIES

At 18 to 20 weeks’ gestation, all pregnant women should be offered a detailed ultrasound that meets previously established minimum standards.\(^{71}\) Most major fetal anatomic abnormalities should be detected by this screen. In particular, the majority of open neural tube defects should be detected by this ultrasound.\(^{72}\) In addition, ultrasound can detect “soft markers,” which are features that increase the a priori risk of fetal aneuploidy but can also be variations of normal. When used alone, second trimester ultrasound soft markers do not effectively discriminate between unaffected fetuses and fetuses with Down syndrome, because of the high positive rate from the large number of potential markers.\(^{73-76}\) Ultrasound soft markers were comprehensively reviewed in a 2007 SOGC guideline,\(^{77}\) and both soft markers and anomalies identified in the 18- to 20-week ultrasound can be used to modify any a priori risk established by age or prior screening. In the absence of soft markers and anomalies, a reduction of risk can be applied. In this circumstance, a conservative negative likelihood ratio of 0.5 is often used, based on studies that have shown an odds ratio, in the presence of a normal fetal ultrasound, ranging from 0.2 to 0.5.\(^{78-86}\) Unless centre-specific levels are determined through clinical audit. However, this should be done only in an established centre performing tertiary level scans.

Recommendation

| 12. The presence or absence of soft markers or anomalies in the 18- to 20-week ultrasound can be used to modify the a priori risk of aneuploidy established by age or prior screening. (II-2B) |

744 ● JULY JOGC JUILLET 2011
FACTORs POTENTIALLY AFFECTING SCREENING PERFORMANCE

A number of factors may affect screening performance. These include accuracy of gestational dating, maternal weight, ethnicity, insulin-dependent diabetes mellitus, accuracy of NT and serum marker measurements, and the use of assisted reproduction technologies.

Gestational Dating

Accurate dating is important. Ultrasound improves the precision of gestational age estimation, and hence reduces the standard deviation of each screening marker. This effect is greater for markers whose concentrations change most with gestational age. For all marker combinations, the FPR is lower by about 2% when gestational age is estimated using a scan. For example, for a DR of 85%, scan dating could reduce the FPR of serum IPS from 4.2% to 2.7%.5

Maternal Weight

There is a negative association between the levels of maternal serum markers and maternal weight, which is due to the dilution effect produced by the physiologic increase in blood volume.84 The trend with first trimester markers is similar to that seen with second trimester markers.8 With second trimester screening, maternal weight adjustment increases DR by about 1% for a given FPR, or reduces FPR by 0.2% for a given DR.11 Weight adjustment is beneficial if there is a marginally elevated AFP when screening for ONTD. When interpreting measurements of serum markers, many screening centres routinely adjust for maternal weight. It has been suggested that published weight correction formulas may not be optimal because of differences in mean weight between the population served and the populations used to derive the formulas. Each laboratory should calculate its own weight adjustment formulas.84

Weight adjustment does not appear to be necessary for NT risk adjustment, because it increases by only a clinically insignificant amount with increasing maternal weight.85

Ethnic Origin

There are differences in the levels of screening markers between women of different ethnic origins after accounting for maternal weight. Maternal serum AFP is 15% higher, total hCG is 18% higher, inhibin A is 8% lower, and PAPP-A is 35% higher in Black women than in Caucasian women. AFP is 6% lower, uE3 is 7% higher, total hCG is 6% higher and PAPP-A is 17% higher in South Asian women. Higher levels of first trimester PAPP-A and β-hCG are seen in Asian women, and higher uE3 is seen in Aboriginal Canadian women.11,86-90 Adjusting for ethnic origin slightly increases the DR for a given FPR, but, more importantly, it tends to equalize the FPR among women of different ethnic groups.11

Statistically significant differences in NT measurement have been found between ethnic groups.90-92 However, it seems these differences may be too small to warrant correction.98

Insulin-Dependent Diabetes Mellitus

Some second trimester serum markers tend to be lower in women with insulin-dependent diabetes mellitus. After weight correction, AFP is ~10% lower and uE3 is ~5% lower in diabetic women. No change in other markers in diabetic women has been demonstrated.11,93,94 To allow for the difference, the observed MoM for a woman with diabetes is divided by the corresponding median MoM in diabetic women without Down syndrome pregnancies. Because of the lack of data in diabetic women who have a Down syndrome pregnancy, a “pseudo risk” can be calculated for diabetic women.11

It appears that NT measurement, free β-hCG, and PAPP-A in women with and without insulin-dependent diabetes are not significantly different.95

Assisted Reproduction

When a pregnancy is a result of IVF, the maternal age used for the determination of the risk of Down syndrome is the age of the donor at the time the egg was harvested.

Data from most published studies show second trimester serum levels of β-hCG and total hCG are higher, and uE3 is lower in pregnancies conceived through IVF.96-99 There were no significant differences in the levels of AFP and inhibin A between IVF and non-IVF pregnancies.97 The variation in hCG is said to be driven by the continuing high progesterone concentrations following hormonal treatment.97 Because of the higher hCG and lower uE3 levels, the FPR of second trimester screening is nearly doubled in IVF pregnancies.97,100,101 In 1999, Wald et al.97 suggested that adjustments for IVF pregnancies could avoid this high FPR. However, results from a recent study in France based on ~1000 IVF pregnancies found no differences in the values of maternal serum AFP, uE3, and hCG between IVF pregnancies and controls. The FPR was similar in the 2 groups.102

In the first trimester, a lower value of PAPP-A has been reported in IVF pregnancies, but data on NT and first trimester free β-hCG remain inconsistent.102,103-106 Many screening programs routinely collect information on IVF; however, whether adjustment is necessary needs further investigation.
Recommendation

13. Information such as gestational dating, maternal weight, ethnicity, insulin-dependent diabetes mellitus, and use of assisted reproduction technologies should be provided to the laboratory to improve accuracy of testing. (II-2A)

GENERAL CONSIDERATIONS

Screening practice differs across Canada and will also change over time. Practitioners should stay updated on the screening modalities available in their areas.

Recommendations

14. Health care providers should be aware of the screening modalities available in their province or territory. (III-B)

15. A reliable system needs to be in place ensuring timely reporting of results. (III-C)

16. Screening programs should be implemented with resources that support audited screening and diagnostic laboratory services, ultrasound, genetic counselling services, patient and health care provider education, and high quality diagnostic testing, as well as resources for administration, annual clinical audit, and data management. In addition, there must be the flexibility and funding to adjust the program to new technology and protocols. (II-3B)

REFERENCES


17. Mulvey S, Zachariah R, Mcllhawie K, Wallace EM. Do women prefer to have screening tests for Down syndrome that have the lowest screen-positive rate or the highest detection rate? Prenat Diagn 2003;23:828–32.


### APPENDIX. SCREENING TERMINOLOGY

**Affected individuals:** Individuals who have the disorder for which the screen is being performed.

**Cut-off level:** The value of a test variable that distinguishes screen positive from screen negative results. The screening cut-off will affect both the detection and false-positive rates: the higher the cut-off, the lower the false-positive rate and the lower the detection rate.

**Detection rate (DR) or sensitivity:** The proportion of affected individuals with positive screening results (usually expressed as a percentage).

**False-positive rate (FPR):** The proportion of unaffected individuals with positive screening results (usually expressed as a percentage). It is the complement of the specificity.

**First trimester window:** The period between 11+0 weeks and 13+6 weeks of gestation.

**Incidence:** The number of new cases of a disorder that arise during a specific period of time, such as a year. This is usually expressed as a rate per 1000.

**Integrated prenatal screen (IPS):** The use of markers in the first and second trimester to calculate the overall risk for fetal aneuploidy.

**Likelihood ratio (LR):** The likelihood that a given test result would be expected in a patient with the target disorder compared with the likelihood that that same result would be expected in a patient without the target disorder. The likelihood ratio for a population is the detection rate divided by the false-positive rate.

**Multiple of the median (MoM):** The observed value of a specific marker divided by the median value for that marker in a specified population (in prenatal screening, usually pregnancies of the same gestational age).

**Marker:** A biological measurement that when present at an abnormal level may indicate the presence of disease.

**Negative predictive value:** The number of unaffected individuals with negative results (true negatives) divided by the total number of individuals with a negative result, both affected and unaffected.

**Odds of being affected given a positive result (OAPR):** The ratio of the number of affected individuals with positive test results to the number of unaffected individuals with positive results.

**Positive predictive value (PPV):** The number of affected individuals with positive results (true positives) divided by the total number of individuals with positive results, both affected and unaffected. It is the odds of being affected given a positive result expressed as a proportion or percentage.

**Positive rate:** The sum of true and false positives. For most screens, the positive rate is virtually equal to the false-positive rate but as the FPR decreases, this becomes a less reliable approximation. The screen positive rate is a useful parameter for the estimation of resource requirements for follow-up services.

**Prevalence:** The number of cases of a disorder present at a point in time or during a specified period. This is usually expressed as a rate per 1000.

**Quality assurance:** The policy, procedures, and systematic actions established in an enterprise for the purpose of providing and maintaining a specified degree of confidence of a screening test.

**Receiver operator curve (ROC):** It is a plot of the true positive rate against the false-positive rate for the different possible cut points of a test. An ROC curve demonstrates the trade-off between sensitivity and specificity (any increase in sensitivity will be accompanied by a decrease in specificity). Accuracy of the test is measured by the area under the ROC curve.

**Second trimester window:** The period between 15+0 weeks and 20+6 weeks of gestation.

**Specificity:** The proportion of unaffected individuals with negative results.

**Unaffected individuals:** Individuals who do not have the disorder for which the screen is being performed.
Abstract

Objective: To evaluate ultrasound “soft markers” used in fetal genetic screening.

Options: Ultrasound screening at 16 to 20 weeks is one of the most common genetic screening and (or) diagnostic tests used during pregnancy. The practical concern for ultrasound screening is false-positive and false-negative (missed or not present) results. The use and understanding of ultrasound soft markers and their screening relative risks is an important option in the care of pregnant women. Currently, the presence of a “significant” ultrasound marker adds risk to the likelihood of fetal pathology, but the absence of soft markers, except in controlled situations, should not be used to reduce fetal risk.

Key Words: Ultrasound, soft marker, prenatal screening, fetus, aneuploidy, trisomy, genetic

Outcomes: The use of ultrasound in pregnancy has significant health and economic outcomes for families and the health care system, compared with no ultrasound use. The Society of Obstetricians and Gynaecologists of Canada (SOGC) recommends a single “routine” ultrasound evaluation at 16 to 20 weeks in all pregnancies. Patients need to be counselled about the positive and negative findings that ultrasound may reveal so they are prepared for unexpected pregnancy knowledge and the possibility of further testing options being offered.

Evidence: Committee members were asked to review specific soft marker ultrasound topics after consensus was reached on the most commonly published soft markers. Medline and PubMed databases were searched for peer-reviewed English articles published from 1985 to 2003. Reviews of each soft marker topic were written by committee members with quality of evidence and classification of recommendations. These reviews were then circulated and discussed by the combined committee. Final format for the guideline was completed by the committee chairpersons.

Values: The quality of evidence and classification of recommendations followed discussion and consensus by the combined committees of Diagnostic Imaging and Genetics of the SOGC.

Benefits, Harms, Costs: It is not possible at this time to determine the benefits, harms, and costs of the guideline because this would require health surveillance and research and health resources not presently available; however, these factors need to be evaluated in a prospective approach by provincial and tertiary initiatives. Consideration of these issues is in the options and outcome section of this abstract.

Recommendations:

1. The screening ultrasound at 16 to 20 weeks should evaluate 8 markers, 5 of which (thickened nuchal fold, echogenic bowel, mild ventriculomegaly, echogenic focus in the heart, and choroid plexus cyst) are associated with an increased risk of fetal aneuploidy, and in some cases with nonchromosomal problems, while 3 (single umbilical artery, enlarged cisterna magna, and pyelectasis) are only associated with an increased risk of nonchromosomal abnormalities when seen in isolation (II-2 B).

2. Identification of soft markers for fetal aneuploidy requires correlation with other risk factors, including history, maternal age, and maternal serum testing results (II-1 A).

3. Soft markers identify a significant increase in fetal risk for genetic disease. Timely referral for confirmation, counselling, and investigation is required to maximize management options (III-B).

Validation: Peer-reviewed guideline development is part of the committee process in addition to SOGC council and editorial review.

Sponsors: SOGC.


These guidelines reflect emerging clinical and scientific advances as of the date issued and are subject to change. The information should not be construed as dictating an exclusive course of treatment or procedure to be followed. Local institutions can dictate amendments to these opinions. They should be well documented if modified at the local level. None of these contents may be reproduced in any form without prior written permission of the SOGC.
INTRODUCTION

Providing an obstetric ultrasound at 16 to 20 weeks’ gestation has become standard practice in Canada.1–3 Although there are many potential benefits, the primary reason to routinely offer this scan is for the detection of fetal abnormalities.4–6 Some obstetric ultrasound findings are considered variants of normal but are noteworthy because they also increase the risk for underlying fetal aneuploidy. These findings are known as “soft markers” and should be considered distinct from fetal anatomic malformations and (or) growth restriction that also increase perinatal and genetic risks.

The presence of soft markers increases the risk for fetal aneuploidy but is not diagnostic. Individual soft markers will vary in the degree of association with fetal aneuploidy. It has become practice to estimate the degree of association as a likelihood ratio (LR) by which the a priori background risk is altered. Detection of multiple soft markers will increase the significance of the finding, compared with seeing the same marker in isolation.7,8 Nonsonographic factors, including maternal age, gestational age, past history, and family history also influence the chance for aneuploidy and should be considered to establish an accurate a priori risk.9–12 In addition, maternal serum testing as an alternate screening tool can complement and enhance the overall screening process.13–18 Providing an accurate assessment of fetal genetic risk requires the ability to integrate known factors before patients can make an informed choice about proceeding with invasive diagnostic testing.

The purpose of this guideline is to (1) evaluate the usefulness of each ultrasound soft marker, (2) assess whether a specific soft marker should be looked for routinely on screening ultrasound, (3) review potential nonkaryotypic implications for soft markers, (4) suggest follow-up recommendations to deal with soft markers once detected, and (5) provide assessment of the quality of information regarding each marker. (See Table 1 for the quality of evidence and classification of recommendation).19

REFERENCES

FETAL SOFT MARKERS USEFUL FOR SCREENING ULTRASOUND

ECHOGENIC INTRACARDIAC FOCUS (Figure 1)

Definition and Imaging Criteria

Echogenic intracardiac focus (EICF) is defined as a focus of echogenicity comparable to bone, in the region of the papillary muscle in either or both ventricles of the fetal heart.1–6 Eighty-eight percent are only in the left ventricle, 5% are only in the right, and 7% are biventricular.7 A grading system has been proposed comparing the echogenicity of the intracardiac focus with surrounding bone. Grade 2 suggests that echogenicity is equal to bone, and grade 3 suggests it is greater.8 Using an appropriate transducer frequency (≤ 5 MHz) and appropriate gain setting, an EICF can be diagnosed on the standard 4-chamber view of the fetal heart.

Association With Fetal Aneuploidy

The association between isolated EICF and fetal aneuploidy has been described in both retrospective and prospective studies. The evidence is best for left or biventricular EICF, but this is likely due to the greater frequency that foci are found in these locations.1–11 A meta-analysis has suggested a likelihood ratio of 2.8 (95% confidence interval [CI] 1.5–5.5).12 However, most studies were undertaken in high-risk women. When the low-risk population is evaluated, the finding of an isolated EICF is associated with lower LRs, from 0–1.8.13–17 Consensus of the SOGC Imaging and Genetics Committees suggests an LR of 2.

Summary

Although the numbers are small, studies suggest that the less frequent right-sided, biventricular, multiple, or particularly conspicuous EICF appear to be associated with a higher risk for fetal aneuploidy, compared with the more common single, left ventricular EICF.8,11,18–21

Association With Nonchromosomal Abnormalities

EICF has not been associated with congenital heart disease or other chromosomal abnormalities.22–25 There may be some ethnic difference regarding the incidence (Asian more often than Caucasian) of EICF.26

Table 1. Criteria for quality of evidence assessment and classification of recommendations

<table>
<thead>
<tr>
<th>Level of evidence*</th>
<th>Classification of recommendations†</th>
</tr>
</thead>
<tbody>
<tr>
<td>I: Evidence obtained from at least one properly designed randomized controlled trial.</td>
<td>A. There is good evidence to support the recommendation for use of a diagnostic test, treatment, or intervention.</td>
</tr>
<tr>
<td>II-1: Evidence from well-designed controlled trials without randomization.</td>
<td>B. There is fair evidence to support the recommendation for use of a diagnostic test, treatment, or intervention.</td>
</tr>
<tr>
<td>II-2: Evidence from well-designed cohort (prospective or retrospective) or case-control studies, preferably from more than one centre or research group.</td>
<td>C. There is insufficient evidence to support the recommendation for use of a diagnostic test, treatment, or intervention.</td>
</tr>
<tr>
<td>II-3: Evidence from comparisons between times or places with or without the intervention. Dramatic results from uncontrolled experiments (such as the results of treatment with penicillin in the 1940s) could also be included in this category.</td>
<td>D. There is fair evidence not to support the recommendation for a diagnostic test, treatment, or intervention.</td>
</tr>
<tr>
<td>III: Opinions of respected authorities, based on clinical experience, descriptive studies, or reports of expert committees.</td>
<td>E. There is good evidence not to support the recommendation for use of a diagnostic test, treatment, or intervention.</td>
</tr>
</tbody>
</table>

*The quality of evidence reported in these guidelines has been adapted from the Evaluation of Evidence criteria described in the Canadian Task Force on the Periodic Health Exam.19
†Recommendations included in these guidelines have been adapted from the Classification of Recommendations criteria described in the Canadian Task Force on the Periodic Health Exam.19
**Recommendations**

1. EICF should be evaluated as part of the 4-chamber cardiac review during the 16- to 20-week ultrasound (III-B).

2. Isolated EICF with a fetal aneuploidy risk less than 1/600 by maternal age (31 years) or maternal serum screen requires no further investigations (III-D).

3. Women with an isolated EICF and a fetal aneuploidy risk greater than 1/600 by maternal age (31 years) or maternal serum screening should be offered counselling regarding fetal karyotyping (II-2 B).

4. Women with right-sided, biventricular, multiple, particularly conspicuous, or nonisolated EICF should be offered referral for expert review and possible karyotyping (II-2 A).

**References**


MILD PYELECTASIS (Figure 2)

Definition and Imaging Criteria

Mild pyelectasis is defined as a hypoechoic spherical or elliptical space within the renal pelvis that measures ≥ 5 mm and ≤ 10 mm. The measurement is taken on a transverse section through the fetal renal pelvis using the maximum anterior-to-posterior measurement. Measurements < 5 mm are normal, should not be designated as pyelectasis, and should not be reported. Pyelectasis may also be referred to as “mild renal pelvic dilatation” or “mild hydronephrosis.”

Association With Fetal Aneuploidy

Isolated pyelectasis is seen in 0.7% of fetuses at 16 to 26 weeks’ gestation. It is an isolated finding in fetal Down syndrome in approximately 2%. Although the likelihood ratio for Down syndrome is approximately 1.9, the 95% CI does cross 1 (0.7–5.1), indicating lack of significance. In the absence of other risk factors, the chance of Down syndrome in the presence of isolated mild pyelectasis remains small and does not justify an invasive diagnostic procedure.

Association With Nonchromosomal Abnormalities

Fetal pyelectasis is associated with congenital hydronephrosis, which is a commonly encountered birth defect. Renal pelvis measurements > 10 mm should be considered equivalent to congenital hydronephrosis with appropriate follow-up. All fetuses with renal pelvic measurements ≥ 5 mm should have a neonatal ultrasound, and
those having measurements > 10 mm should also have a third trimester ultrasound.2

Summary

Evaluation of fetal kidneys, which includes possible pyelectasis, is considered part of the routine screening ultrasound at 16 to 20 weeks' gestation and should be reported.8 The finding of isolated pyelectasis does not appear to significantly increase the risk of fetal aneuploidy in low-risk women and does not justify invasive prenatal testing, but noninvasive maternal serum screening may assist in risk assessment. Owing to the increased risk of fetal hydronephrosis, a neonatal follow-up scan should be arranged in all cases of mild isolated pyelectasis. A third trimester follow-up ultrasound should only be considered if pyelectasis is ≥ 10 mm. Referrals should be considered for women aged over 35 years and for women who have additional ultrasound findings, renal pelvis measurements > 10 mm, or maternal serum screening results showing increased chromosomal risks.

Recommendations

1. Evaluation of fetal kidneys is a part of the screening ultrasound at 16 to 20 weeks,7 and if pyelectasis is visualized, the renal pelvis should be measured in the anterior/posterior diameter (III-B).

2. All fetuses with renal pelvic measurements ≥ 5 mm should have a neonatal ultrasound, and those having measurements > 10 mm should be considered for a third trimester scan (II-2 A).

3. Isolated mild pyelectasis does not require fetal karyotyping (II-2 E).

4. Referral for pyelectasis should be considered with additional ultrasound findings and/or in women at increased risk for fetal aneuploidy owing to maternal age or maternal serum screen results (II-2 A).

References


SINGLE UMBILICAL ARTERY (Figure 3)

Definition and Imaging Criteria

Single umbilical artery (SUA) is the absence of one of the arteries surrounding the fetal bladder and in the fetal umbilical cord. Assessment of the umbilical arteries can be made from the cord itself in either transverse or longitudinal sections.1–3 The umbilical arteries can also be assessed at the cord insertion site into the fetal abdomen and on either side of the fetal bladder as the vessels originate from the iliac arteries. If needed, the assessment can be enhanced with colour flow Doppler.

Association With Fetal Aneuploidy

Isolated SUA has not been found to be significantly associated with fetal aneuploidy.1–6

Association With Nonchromosomal Abnormalities

Isolated SUA has been associated with both underlying fetal renal and cardiac abnormalities,1–7–9 as well as low birth weight.5,8

Summary

Assessment of cord vessels is considered a part of the routine obstetric ultrasound at 16 to 20 weeks.10 The finding of a single umbilical artery warrants a detailed review of fetal anatomy, including kidneys and heart (fetal echo). Appropriate fetal growth should be confirmed through clinical evaluation with follow-up ultrasound for clinical concerns. An isolated SUA does not warrant invasive testing for fetal aneuploidy.

Recommendations

1. Assessment of cord vessels is considered a part of the routine obstetric ultrasound at 16 to 20 weeks (III-A).

2. The finding of a single umbilical artery requires a more detailed review of fetal anatomy, including kidneys and heart (fetal echo) (II-2 B).

3. An isolated single umbilical artery does not warrant invasive testing for fetal aneuploidy (II-2 A).

References


ECHOCGENIC BOWEL (Figure 4)

Definition and Imaging Criteria
Echogenic bowel is defined as fetal bowel with homogeneous areas of echogenicity that are equal to or greater than that of surrounding bone.¹ The echogenicity has been classified as either focal or multifocal.² There have been various techniques used to define echogenic bowel, partially because of concerns raised about intra- and interobserver variability.³ A grading system based on comparison of the echogenicity of fetal bowel and surrounding bone relative to the ultrasound machine gain setting minimizes observer variability and should be used. Grade 2 suggests that echogenicity is equal to bone whereas grade 3 suggests that it is greater.³ Whenever echogenic bowel is suspected, the gain setting should be lowered to enable this comparison and to ensure that bowel hyperechogenicity is real.³ This should help to minimize a false-positive diagnosis of hyperechogenicity.

Association With Fetal Aneuploidy
The presence of echogenic bowel is associated with an increased risk for fetal aneuploidy, including trisomy 13, 18, 21, and the sex chromosomes. It has been detected in 0.6% to 2.4% of all second trimester fetuses²,⁴–⁹ and as an isolated finding in 9% of fetuses with aneuploidy (2.8% to 25%).²–¹⁹ As a result, it has been suggested that the likelihood ratio for this marker is 6 (CI 2.7–6.8).⁶

Association With Nonchromosomal Abnormalities
The presence of echogenic bowel has been associated with an increased risk for cystic fibrosis in the fetus, congenital infection, intra-amniotic bleeding, congenital malformations of the bowel, and other perinatal complications, including intrauterine growth restriction. The risk of cystic fibrosis in the fetus with echogenic bowel is approximately 2% (0 to 13%).³,¹⁰–¹³,¹⁸–²¹ The a priori risk will change if the
parental carrier status is known. The association between congenital infection and hyperechogenic bowel has been noted for the most common pathogens known to cause fetal infections (cytomegalovirus [CMV], herpes, parvovirus, rubella, varicella, and toxoplasmosis).\textsuperscript{3,4,6,11,12,14,18,19} Intra-amniotic bleeding has also been identified as an etiology of echogenic bowel. This can result from intra-amniotic bleeding owing to placental abruptions or invasive procedures.\textsuperscript{18,19,22–24} Congenital malformations of the fetal bowel can lead to increased echogenicity. Studies have suggested that this is more likely with upper gastrointestinal (GI) lesions. Other ultrasound features, such as ascites and dilated loops of bowel, will often be present in this circumstance.\textsuperscript{18,19,25–27} Echogenic bowel has also been reported with poor fetal growth, which is associated with an increase in perinatal morbidity and mortality.\textsuperscript{4–6,10–14,18,19,28}

**Summary**

Evaluation of the fetal abdomen is an established component of the screening obstetric ultrasound at 16 to 20 weeks.\textsuperscript{29} This includes an evaluation of bowel echogenicity using an appropriate transducer (5 MHZ or less) and ultrasound gain setting. Echogenic bowel is associated with a significantly increased risk for both chromosomal and nonchromosomal fetal abnormalities. Timely referral for validation, consultation, and further investigation is important.

Further evaluations may include a detailed review of fetal anatomy, growth, and placental characteristics. Laboratory investigations may include a fetal karyotype, DNA testing for cystic fibrosis, and testing for congenital infections (maternal serum titres, fetal amniotic culture, or polymerase chain reaction [PCR] for viral DNA). A maternal serum screen may be considered because elevations in alpha fetoprotein and hCG in the presence of echogenic bowel may further define a population at increased risk for perinatal morbidity and mortality. Obstetric and ultrasound follow-up may also be important.

**Recommendations**

1. Evaluation of the fetal bowel should be done routinely during the 16–20-week obstetric ultrasound (III-B).
2. Echogenic bowel should be identified by comparison with the echogenicity of surrounding bone using an appropriate transducer and gain setting. Bowel echogenicity equal to or greater than bone is significant (grade 2 or 3) (II-2 A).
3. No further investigations are required for grade 1 echogenic bowel (II-2 D).
4. Grade 2 and 3 echogenic bowel is associated with both chromosomal and nonchromosomal abnormalities. Expert review is recommended to initiate the following: a. detailed ultrasound evaluation looking for additional structural anomalies or other soft markers of aneuploidy (II-2 A); b. detailed evaluation of the fetal abdomen looking for signs of bowel obstruction or perforation (II-2 B); and c. detailed evaluation of placental characteristics (echogenicity, thickness, position, and placental cord insertion site) (II-2 B); d. genetic counselling (II-2 A); e. laboratory investigations that...
should be offered, including fetal karyotype, maternal serum screening, DNA testing for cystic fibrosis (if appropriate), and testing for congenital infection (II-2A).

References


THICKENED NUCHAL FOLD (Figure 5)

Definition and Imaging Criteria

The nuchal fold is the skin thickness in the posterior aspect of the fetal neck. A nuchal fold measurement is obtained in a transverse section of the fetal head at the level of the cavium septum pellicidum and thalami, angled posteriorly to include the cerebellum. The measurement is taken from the outer edge of the occiput bone to the outer skin limit directly in the midline. The definition of a thickened nuchal fold has varied, although many researchers and centres now use gestational-age specific criteria. Consensus for this document is that a measurement ≥ 6 mm be considered significant between 18 and 24 weeks and a measurement of ≥ 5 mm be considered significant at 16 to 18 weeks. A thickened nuchal fold should be distinguished from cystic hygroma, in which the skin in this area has fluid-filled loculations. A thickened nuchal fold should not be confused with nuchal translucency, which is a specific measurement of fluid in the posterior aspect of the neck at 11 to 14 weeks’ gestation.

Association With Fetal Aneuploidy

A meta-analysis reviewed the performance of a thick nuchal fold at 6 mm or greater and showed that the risk for Down syndrome increased by approximately 17-fold (CI 8–35).
Association With Nonchromosomal Abnormalities

A thickened nuchal fold can be associated with single gene abnormalities, such as Noonan syndrome, multiple pterygium syndrome, and skeletal dysplasias.7,8 Thickened nuchal fold has also been associated with congenital cardiac defects.7,9,10

Summary

Evaluation of the nuchal fold should be considered during the screening ultrasound at 16 to 22 weeks’ gestation. A nuchal fold of 6 mm or greater at 18 to 24 weeks or of 5 mm or greater at 16 to 18 weeks should be considered significant and should prompt referral for validation and consultation. The finding of an isolated thickened nuchal fold significantly increases the risk for fetal aneuploidy, and fetal karyotyping should be offered. Centres may use alternate definitions, taking into account gestational age and other risk factors. Nuchal index has been described as an effective method to deal with the normal increase in nuchal fold measurement that accompanies advancing gestational age. Nuchal index is the mean nuchal fold/mean biparietal diameter (BPD) × 100. A value of 11 or greater has a sensitivity of 50% and a specificity of 96%.11

The suggested association of nuchal fold thickening and congenital heart defect is based on small studies. Careful detailed ultrasound examination, including the 4-chamber view and outflow tracts, should be performed. The rare occurrence of an underlying syndromic etiology for the increased nuchal fold justifies a directed, detailed anatomic survey of the fetus and a careful newborn examination.12

Recommendations

1. Nuchal fold measurement should be a part of the screening obstetric ultrasound at 16 to 20 weeks (III-B).

2. A thickened nuchal fold significantly increases the risk of fetal aneuploidy. Expert review is recommended, and karyotyping should be offered (II-1 A).

3. A thickened nuchal fold is associated with congenital heart disease and rarely with other genetic syndromes. Expert review is recommended (II-2 B).

References


**MILD VENTRICULOMEGALY (Figure 6)**

**Definition and Imaging Criteria**

Cerebral ventriculomegaly is defined by atrial measurements ≥ 10 mm. Mean atrial measurements are 7.6 mm, standard deviation (SD) 0.6 mm. Mild ventriculomegaly (MVM) is defined as measurements ≥ 10 to ≤ 15 mm. Measurements are obtained from an axial plane at the level of the thalamic nuclei just below the standard image to measure the BPD. Ventricular measurements are usually obtained in the far image field because of “typical” near-field artifacts. Cursors are positioned perpendicular to the long axis of the ventricle at the edges of the ventricular lumen, near the posterior portion of the choroid plexus.

**Association With Fetal Aneuploidy**

When MVM is isolated, the incidence of abnormal fetal karyotype is estimated at 3.8% (0 to 28.6%). Idiopathic lateral ventriculomegaly is found in approximately 0.15% of chromosomally-normal fetuses, whereas 1.4% of trisomy 21 fetuses in the second trimester have idiopathic ventriculomegaly. This suggests a likelihood ratio of 9 for the risk of karyotype abnormality.

**Association With Nonchromosomal Abnormalities**

Fetal ventriculomegaly is the most commonly detected ultrasonographic abnormality of the central nervous system. Ventriculomegaly can arise from agenesis of the corpus callosum, cerebral maldevelopment or destruction, vascular anomalies, or an obstruction within the ventricular system. Children with a prenatal diagnosis of MVM have abnormal neurodevelopment in 10% to 36% of cases dependent on associated anomalies, etiology, and ventricular measurement. In combined case series, mortality is reported at 3.7%. When MVM resolves, abnormal outcome has been reported but is infrequent (< 10%). Unilateral MVM also carries a favourable prognosis when isolated. After the prenatal diagnosis of MVM, maternal evaluation for congenital infection is recommended. Amniocentesis should be offered for karyotype and congenital infection assessment. Other imaging modalities such as magnetic resonance imaging (MRI) might be considered.
Summary
Lateral ventriculomegaly can be detected on standard cranial biometry planes and should be evaluated on both screening ultrasounds as well as detailed ultrasound for higher risk women. The ventricles should be measured if they appear to be larger than the choroid plexus. The finding of ventriculomegaly should prompt a timely referral for consultation and validation. Evaluation of lateral ventriculomegaly should include a detailed examination of fetal anatomy, including the heart. Neonatal assessment and follow-up are important to rule out associated abnormalities because of the potential for abnormal neurodevelopment.

Recommendations
1. Fetal cerebral ventricles should be measured if they subjectively appear larger than the choroid plexus (III-B).
2. Cerebral ventricles greater than or equal to 10 mm are associated with chromosomal and central nervous system pathology. Expert review should be initiated to obtain the following: a. a detailed anatomic evaluation looking for additional malformations or soft markers (III-B); b. laboratory investigation for the presence of congenital infection or fetal aneuploidy (III-B); and c. MRI as a potential additional imaging technique (II-2 C).
3. Neonatal assessment and follow-up are important to rule out associated abnormalities and are important because of the potential for subsequent abnormal neurodevelopment (II-2 B).

References

CHOROID PLEXUS CYSTS (Figure 7)

Definition and Imaging Criteria
Choroid plexus cysts (CPCs) are sonographically discrete, small cysts (≥3 mm) found in the choroid plexus within the lateral cerebral ventricles of the developing fetus at 14 to 24 weeks' gestation. Imaging of the choroid plexus is performed in the transverse plane of the fetal head at the same level that the lateral cerebral ventricle is evaluated. The choroid plexus should be inspected bilaterally for the presence of cysts. The size of CPCs is not of clinical relevance. Evaluation of the choroid plexus in the near field ventricle will be more difficult owing to imaging artifact.

Association With Fetal Aneuploidy
CPCs have been identified in 1% of fetuses during the second trimester screening ultrasound. The incidence of CPCs is 50% in fetuses with trisomy 18; however, only 10% of fetuses with trisomy 18 will have CPCs as the only identifiable sonographic marker on ultrasound screening. The likelihood ratio for trisomy 18 when an isolated CPC is identified is 7 (95% CI 4–12). The number of cysts and the cysts’ distribution or size does not change the risk. Although it has been suggested that an isolated CPC may increase the risk for trisomy 21 with a likelihood ratio of 1.9, the 95% CI crosses 1 and lacks statistical significance. The presence of CPCs in chromosomally normal fetuses is not associated with other fetal abnormalities or abnormal postnatal development.

Association With Nonchromosomal Abnormalities
CPCs may be present in approximately 5% of normal cases, however, CPCs should not be considered a marker for fetal abnormality. Evaluation of the fetal cranium, including the ventricles and choroid plexus, is considered part of the routine screening.
ultrasound at 16 to 20 weeks’ gestation.\textsuperscript{19} Identification and reporting of CPCs should be a part of this screening examination. With the presence of CPCs, caregivers should next evaluate maternal age risk and, if available, the maternal serum screen.\textsuperscript{2} CPCs increase the risk for trisomy 18. Follow-up ultrasound is not necessary for isolated CPCs. Referral for counselling and possible invasive testing is only necessary if maternal age is 35 years or older or the maternal serum screen is positive for either trisomy 18 or 21.\textsuperscript{2,20}

**Recommendations**

1. Choroid plexus should be evaluated for the presence of discrete cysts during the 16- to 20-week ultrasound (III-B).

2. Isolated CPCs require no further investigation when maternal age or the serum screen equivalent is less than the risk of a 35-year-old (II-2 E).

3. Fetal karyotyping should only be offered if isolated CPCs are found in women 35 years or older or if the maternal serum screen is positive for either trisomy 18 or 21 (II-2 A).

4. All women with fetal CPCs and additional malformation should be offered referral and karyotyping (II-2 A).

5. All women with CPCs and additional soft markers should be offered additional counselling and further ultrasound review (III-B).

**References**


ENLARGED CISTERNA MAGNA (Figure 8)

**Definition and Imaging Criteria**

The cisterna magna is measured on a transaxial view of the fetal head angled 15 degrees caudal to the canthomeatal line. The anterior/posterior diameter is taken between the inferior/posterior surface of the vermis of the cerebellum to the inner surface of the cranium. An enlarged cisternal magna is defined by an anterior/posterior diameter ≥ 10 mm. The measurement will be falsely exaggerated by a steep scan angle through the posterior fossa or dolichocephaly.  

**Association With Fetal Aneuploidy**

An enlarged cisterna magna has been described in association with fetal aneuploidy, particularly trisomy 18. The association with aneuploidy appears to be strongest in the absence of ventricular dilatation but in the presence of other anomalies. There are no large prospective studies to evaluate this marker.

**Association With Nonchromosomal Abnormalities**

An enlarged cisterna magna is commonly seen in association with other anatomic (arachnoid cyst, Dandy Walker malformation, and Dandy Walker variant) and syndromic (oro-facial–digital syndrome, Meckel-Gruber syndrome, and DiGeorge syndrome) abnormalities.

**Summary**

Review of the fetal cerebellum and cisterna magna is a routine part of the screening ultrasound at 16 to 20 weeks’ gestation. If the cisterna magna is subjectively increased, a measurement should be undertaken. The mean diameter of a normal cisterna magna is 5 mm, SD 3 mm. A measurement ≥ 10 mm is considered an abnormality and appropriate referral for consultation and validation should be initiated. A detailed fetal examination should be performed looking for other anomalies, growth restriction, or abnormal amniotic fluid volume. An isolated enlarged cisterna magna is not an indication for fetal karyotyping.

**Recommendations**

1. Review of the fetal cerebellum and cisterna magna is a routine part of the screening ultrasound at 16 to 20 weeks.
If the cisterna magna is subjectively increased, a measurement should be taken (III-B).

2. An isolated enlarged cisterna magna is not an indication for fetal karyotyping (III-D).

3. With an enlarged cisterna magna, expert review is recommended for follow-up ultrasounds and possible other imaging modalities (for example, MRI) and investigations (III-B).

4. If the enlarged cisterna magna is seen in association with other abnormal findings, fetal karyotyping should be offered (III-B).

REFERENCES

FETAL SOFT MARKERS USEFUL FOR COMPREHENSIVE ULTRASOUND

SHORT FEMUR LENGTH

Definition and Imaging Criteria
A short femur length is defined as either a measurement below the 2.5th percentile for gestational age or a measurement that is less than 0.9 of that predicted by the measured biparietal diameter. The femur should be measured with the bone perpendicular to the ultrasound beam and with epiphyseal cartilages visible but not included in the measurement. The relation between bone length and head size may differ across racial groups.

Association With Fetal Aneuploidy
Short femur length has been found to have a sensitivity of 16% in the prediction of Down syndrome with a false-positive rate of 4%. A meta-analysis showed a likelihood ratio of 2.7 (95% CI 2.1–6.0).

Association With Nonchromosomal Abnormalities
Short femur length can also be associated with skeletal dysplasias or fetal growth restriction.

Summary
Short femur length is an ultrasound marker for fetal aneuploidy, particularly trisomy 21. The mathematical model used to determine a positive result is not amenable to screening ultrasound; however, it should be included in the panel of markers used by tertiary centres.

If a femur appears abnormal or its length is found to be below the 2.5th percentile for gestational age, it may be indicative of fetal growth restriction or a more general skeletal malformation. In this circumstance, other long bones should be assessed and referral with follow-up ultrasound considered.

Recommendations
1. Although femur length is standard biometry on the 16- to 20-week ultrasound, the assessment for relative shortness is not part of the screening evaluation (III-C).
2. Relative femur shortening is an ultrasound marker for trisomy 21 and should be considered during tertiary level evaluation (II-1 A).
3. If a femur appears abnormal or measures short on screening ultrasound, other long bones should be assessed and referral with follow-up ultrasound considered (III-B).

SHORT HUMERUS LENGTH

Definition and Imaging Criteria
A short humerus length is defined as a length below the 2.5th percentile for gestational age or as a measurement less than 0.9 of that predicted by the measured biparietal diameter. The humerus should be measured with the bone perpendicular to the ultrasound beam and with epiphyseal cartilages visible but not included in the measurement.
**Association With Fetal Aneuplody**

Short humeral length has been found to have a sensitivity of 9% with a false-positive rate of 3%. A meta-analysis showed a likelihood ratio of 7.5 (95% CI 4.5–12).3

**Association With Nonchromosomal Abnormalities**

Short humeral length can also be associated with skeletal dysplasias or fetal growth restriction.4 Humeral length has also been recorded as multiples of the median for gestational age. This allows for a graded response including a negative predictor for the relatively longer humerus.5

**Summary**

Short humeral length is an ultrasound marker for fetal aneuploidy, particularly trisomy 21. Humeral length is not currently part of the screening obstetric ultrasound; however, it should be included in the panel of markers used by tertiary centres. During screening ultrasound, if the humerus appears abnormal or its length is short, other long bones should be assessed and referral with follow-up ultrasound considered.

**Recommendations**

1. Humeral length is not part of the current screening ultrasound at 16 to 20 weeks but should be considered for future inclusion (III-B).

2. Relative humeral shortening is an ultrasound marker for trisomy 21 and should be considered during tertiary level evaluation (II-1 A).

3. If the humerus is evaluated and appears abnormal or short, other long bones should be assessed and referral with follow-up ultrasound considered (III-B).

**References**


**NASAL BONE**

**Definition and Imaging Criteria**

Nasal hypoplasia has been recognized as a feature of postnatal trisomy 21.1 This has led to prenatal evaluation of the nasal bone, which has been shown to be a thin echogenic line within the bridge of the fetal nose. The fetus is imaged facing the transducer with the fetal face strictly in the midline. The angle of insonation is 90 degrees, with the longitudinal axis of the nasal bone as the reference line. Calibres are placed at each end of the nasal bone. Absence of the nasal bone or measurements below 2.5th percentile are considered significant.2–4

**Association With Fetal Aneuplody**

Preliminary second trimester studies appear to confirm that hypoplastic or absent nasal bone is an ultrasound marker for fetal Down syndrome, while, conversely, a normal nasal bone would reduce significantly the risk.5,6 The likelihood ratio for this finding varies depending on ethnic background. Although a hypoplastic nasal bone was associated with an overall likelihood ratio for Down syndrome at 51, it was found to be 132 for Caucasians and 8.5 for African Caribbeans. The negative likelihood ratio was 0.39 for Caucasians and 0.27 for African Caribbeans.7 Nasal hypoplasia has not been associated with other aneuploidy.

**Association With Nonchromosomal Abnormalities**

An absent or hypoplastic nasal bone has not been found to be associated with chromosomal abnormalities.

**Summary**

Hypoplastic or absent nasal bone is an ultrasound marker for fetal Down syndrome, and a normal nasal bone length significantly reduces the risk. Although views of the fetal nasal bone are readily obtained by imaging the facial profile, this is not considered a part of the routine screening ultrasound.8 In circumstances where the facial profile is seen and the nasal bone is felt to be absent or hypoplastic, referral is recommended. Assessment of the nasal bone should be considered for research or tertiary level evaluation.

**Recommendations**

1. Assessment of the fetal nasal bone is not considered a part of the screening ultrasound at 16 to 20 weeks (III-B).

2. Hypoplastic or absence nasal bone is an ultrasound marker for fetal Down syndrome, and if suspected, expert review is recommended (II-2 B).

**References**


Fifth Finger Clinodactyly

Definition and Imaging Criteria
Fifth finger clinodactyly is defined by a hypoplastic or absent mid-phalanx of the fifth digit. Ultrasound identification of the fetal hand must first be undertaken and then appropriate magnification accomplished. The evaluation requires stretching of the 5 fingers. The diagnosis is established when the middle phalanx of the fifth finger is markedly smaller than normal or absent, which often causes the finger to be curved inward.

Association With Fetal Aneuploidy
Fifth finger clinodactyly is found in 60% of neonates affected with Down syndrome. During antenatal screening, it has been found to be present in 3.4% of normal fetuses and in 18.8% of fetuses with Down syndrome. This suggests a likelihood ratio of 5.6 (95% CI 2.5–11.9).

Association With Nonchromosomal Abnormalities
As an isolated finding, clinodactyly is not associated with other nonchromosomal anatomic or syndromic abnormalities.

Summary
Evaluation of the fetal fingers is not an established part of the screening obstetric ultrasound at 16 to 20 weeks’ gestation. The risk for fetal aneuploidy in the presence of isolated clinodactyly has been estimated to increase by 5.5, and although this finding is considered a significant soft marker, it has not been confirmed with prospective studies. In the event that clinodactyly is seen, it is important to initiate timely referral for consultation, validation, and possibly further investigations. Tertiary centres may use evaluation for clinodactyly as part of their review for patients at increased risk for aneuploidy.

Recommendations
1. Imaging of the outstretched hand to evaluate for fifth finger clinodactyly is not an expectation during the 16- to 20-week ultrasound (III-C).
2. Fifth finger clinodactyly is associated with trisomy 21 and should be considered for research or tertiary-level evaluation (III-B).

References

Fetal Soft Markers Not Established for Clinical Practice

Brachycephaly

Definition and Imaging Criteria
Fetuses affected with trisomy 21 are known to be at increased risk for abnormalities in brain growth and maturation. This is known to result in shortening of the frontal occipital brain length primarily owing to a smaller frontal lobe. The subsequent abnormal skull shape (brachycephaly) has been evaluated as a screening tool. Initially, brachycephaly was studied with the cephalic index—the biparietal diameter over the occipital frontal diameter. More recent investigations have specifically studied the hypoplastic frontal lobe with various biometric measurements and calculations.

Association With Fetal Aneuploidy
The cephalic index does not vary significantly between trisomy 21 and euploid fetuses. Other calculations of frontal lobe hypoplasia have shown some screening potential in retrospective studies; however, no prospective studies have been undertaken, and there are no calculated likelihood ratios. The “strawberry” shaped cranium has been specifically described as being associated with trisomy 18 but has not been evaluated prospectively in a low-risk population.

Association With Nonchromosomal Abnormalities
Brachycephaly is not strongly associated with other chromosomal abnormalities.
Summary

Brachycephaly has not been established as an effective screen for fetal aneuploidy. No recommendations for follow-up or changes in neonatal care are advised as a result of a finding of brachycephaly or abnormalities in frontal lobe biometry. Other abnormal cranial morphologies, such as “strawberry”12 or “lemon”13 shapes, are associated with fetal pathology and should prompt appropriate referral.

References


INCREASED ILIAC ANGLE

Definition and Imaging Criteria

It has been identified that postnatal trisomy 21 is associated with a wider lateral flare of the iliac bones. Two techniques have been described to measure the fetal iliac angle.1,2 Both methods use the axial (transverse) view of the fetal pelvis. In one method, the converging lines are drawn along the posterior lateral aspect of the iliac wings, while in the second method, the converging lines are drawn through the middle of the iliac wing extremity. It has been suggested that an angle ≥90 degrees should be considered the upper limit of normal when screening for trisomy 21.1,3

Association With Fetal Aneuploidy

Several prospective and retrospective studies have shown the association between increased iliac angle and trisomy 21.4–8,9 Research to date has been limited to high-risk populations. There is no screening sensitivity for this marker in the low-risk population.

Association With Nonchromosomal Abnormalities

An increased iliac angle has not been associated with specific chromosomal abnormalities.

Summary

Increased iliac angle is a possible marker for trisomy 21; however, measurement techniques do not make it amenable to a screening exam, and it has not been evaluated to be effective in a low-risk population. This marker may be useful for tertiary centres investigating high-risk patients or as a possible negative predictor.9

References

SMALL FETAL EAR LENGTH

Definition and Imaging Criteria
Small low-set ears are a clinical feature in newborns with trisomy 21 and other aneuploidy. Although fetal ear position is difficult to determine sonographically, ear length is possible, and normal ranges have been established. Ear length is measured in a coronal view and defined as the maximal distance between the superior and inferior border of the external ear.

Association With Fetal Aneuploidy
A prospective study has been undertaken to evaluate fetal ear length and its association with fetal aneuploidy. A sensitivity of 32% and a specificity of 93% was found. This might suggest a likelihood ratio between 3 and 5; however, in 29% of fetuses, appropriate imaging was not able to be obtained. Actual likelihood ratios with confidence intervals have not been published.

Association With Nonchromosomal Abnormalities
Small, low-set, and malformed ears are associated with other genetic abnormalities; however, antenatal detection and evaluation are difficult.

Summary
Although short fetal ear length may be a marker for fetal aneuploidy, adequate evaluation has not been undertaken to establish its usefulness as either a screening tool or as part of a panel of markers for tertiary centres. The use of fetal ear length remains relegated to research protocols.

References

SANDAL GAP

Definition and Imaging Criteria
Sandal gap is described as the separation of the great and second toe and has been reported to be present in 45% of newborns with trisomy 21. Prenatal diagnosis requires imaging the foot and toes from the plantar view.

Association With Fetal Aneuploidy
Although sandal gap has been reported as a finding in fetuses with Down syndrome in the third trimester, it is a subtle sonographic finding in the second trimester. No studies have been undertaken to establish a risk for aneuploidy based on this finding.

Association With Nonchromosomal Abnormalities
The finding of sandal gap may be a normal variant and is not associated with other chromosomal abnormalities.

Summary
No further investigations or follow-up are necessary if isolated sandal gap is detected. It is not part of the screening ultrasound.

References

Recommendations
1. Brachycephaly, increased iliac angle, sandal gap, and fetal ear length are not considered a part of the screening ultrasound at 16 to 20 weeks (III-C).
2. Brachycephaly, increased iliac angle, sandal gap, and fetal ear length should only be evaluated in research protocols or tertiary centres (II-3 D).
3. With specific abnormal cranial morphology such as “clover leaf,” “strawberry,” or “lemon” shapes, referral should be considered (II-2 A).

Discussion
Prenatal diagnosis of fetal aneuploidy is of varying importance to individuals. Diagnosis can only be undertaken with invasive tests that are accompanied by procedure-related risks. Although uncommon, when a complication does occur, it usually results in the loss of a normal fetus. A woman’s decision to proceed with testing will involve an assessment of the risk for the procedure versus the chance of detecting an abnormality. For some, no level of risk assessment for aneuploidy will lead to invasive testing, and as such, screening for the abnormality is of less relevance. It is important to remember that the process of prenatal screening and the decision to proceed with invasive testing...
is voluntary. Caregivers who counsel women must be knowledgeable, must have the ability to integrate various risk factors, and must maintain a nondirective approach.\(^1\)

The diagnosis of and screening for fetal abnormalities make the 16- to 20-week obstetric ultrasound both clinically effective and cost effective.\(^2\)–\(^4\) Based on ultrasound findings, further investigations or treatment may be offered that are gestational-age dependent and thus time limited. If any fetal abnormalities or soft markers are discovered on routine ultrasound, it is important that findings be expeditiously communicated to primary caregivers. Waiting for transcription, editing, and the mail service is unacceptable in this circumstance. Persons who report these findings should do so verbally, electronically, or by fax. Primary caregivers should then relay information to the patient and offer referral for consultation, validation, and possibly further investigation. These referrals will often be to genetic and (or) prenatal diagnostic services that should be capable of urgent accommodation.

Patients who receive news of potential or real fetal abnormalities will experience anxiety and distress.\(^5\) Information should only be given to patients by individuals who can answer preliminary questions and initiate subsequent counselling, referrals and (or) investigations. Although patients will look for answers in the Diagnostic Imaging department, this is seldom the appropriate setting. Patients should

---

### Table 2. Ultrasound “soft markers” performance summary in the detection of aneuploidy (trisomy 21, 18) and other genetic/congenital anomalies

<table>
<thead>
<tr>
<th>Ultrasound “soft markers” (evidence and classification)(^1)</th>
<th>Aneuploidy (LR)(^2)</th>
<th>Congenital/Anomaly Association(^3)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A. Screening scan (16-20 weeks)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nuchal fold (I, A)</td>
<td>17</td>
<td>Congenital heart disease</td>
</tr>
<tr>
<td>Echogenic bowel (II-2, A)</td>
<td>6</td>
<td>CF2%, infection 3%, GI 6%</td>
</tr>
<tr>
<td>Ventriculomegaly (II-2, A)</td>
<td>9</td>
<td>AC, CNS, infection, obstruction</td>
</tr>
<tr>
<td>Echogenic cardiac focus (III, A)</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Choroid plexus cyst (II-2, A)</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Single umbilical artery (III, A)</td>
<td></td>
<td>Renal, cardiac</td>
</tr>
<tr>
<td>Enlarged cisterna magna (III, A)</td>
<td></td>
<td>OFD, MG, DiG</td>
</tr>
<tr>
<td>Renal pyelectasis (II-2, A)</td>
<td></td>
<td>Hydronephrosis; reflux</td>
</tr>
<tr>
<td><strong>B. Comprehensive scan (calculation; detail)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clinodactyly (II-2, A)</td>
<td>5.6</td>
<td></td>
</tr>
<tr>
<td>Humerus (short) (II-2, A)</td>
<td>7.5</td>
<td>skeletal dysplasia; IUGR</td>
</tr>
<tr>
<td>Femur (short) (II-2, A)</td>
<td>2.7</td>
<td>skeletal dysplasia; IUGR</td>
</tr>
<tr>
<td>Nasal bone absent/hypo (II-2, A)</td>
<td>51</td>
<td></td>
</tr>
<tr>
<td><strong>C. Research/Not useful</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brachycephaly (III, B)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Iliac angle (II-2, A)</td>
<td>TBD</td>
<td></td>
</tr>
<tr>
<td>Ear length (III, B)</td>
<td>3–5</td>
<td></td>
</tr>
<tr>
<td>Sandal toe (III, B)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^1\)Canadian Task Force on Periodic Health Examination, Health Canada; Quality of Evidence; Classification of Recommendation (Ann Intern Med 1993; 118:731-7).

\(^2\)LR: likelihood ratio; TBD: to be determined.

\(^3\)CF: cystic fibrosis; CNS: central nervous system; GI: gastrointestinal; OFD: oro-facial-digital syndrome; MG: Meckel Gruber Syndrome; DiG: Di George Syndrome; IUGR: intrauterine growth restriction; AC: agenesis corpus callosum
be told about general concerns and assured that their primary caregiver will receive the report as quickly as possible.

Sixteen potential second trimester soft markers for fetal aneuploidy are reviewed in this document (Table 2). Only 5 markers are considered useful for evaluation for fetal aneuploidy at the time of a screening ultrasound. Increased nuchal fold, echogenic bowel, mild ventriculomegaly, echogenic foci in the heart, and choroid plexus cysts are associated with an increased risk of aneuploidy. Choroid plexus cysts are only associated with trisomy 18 and, in this circumstance, adjustment should only be made for this specific risk. The markers clinodactyly, short humerus, short femur, and hypoplastic or absent nasal bone are all associated with aneuploidy but should be used in tertiary level ultrasounds and (or) research protocols. The mathematical evaluation for short long bones is not part of the screening process and the images for clinodactyly and the nasal bone are not established as a standard part of the 16- to 20-week scan. Three other markers—single umbilical artery, enlarged cisterna magna, and pylectasis—do not have a well-established association with aneuploidy when seen in isolation and should not be used to adjust risk when there are no other significant risk factors. However, these latter findings have other potential perinatal implications, and thus evaluation and reporting remain important during the screening process. Four markers—brachycephaly, iliac angle, ear length, and sandal gap—are not established as markers for screening a low-risk population and should not be evaluated except in a research setting or at a tertiary level.

The reduction in risk that accompanies the absence of ultrasound markers is dependent on the diligence with which an entire panel of markers is evaluated. Risk reduction has only been validated in single institutions or with prospective studies using rigorous research protocols. Although this may be recreated in dedicated prenatal diagnosis centres, a reduction should not be applied on the basis of a 16- to 20-week “screening” scan, owing to the variety of imaging locations involved. In the event that multiple (more than 2) markers are identified, it is recommended that patients be referred for confirmation, counselling, and possible further investigation. It is widely accepted that individual markers function independently, and as a result, when clustered together, they convey an even greater risk. This may be true even for markers that do not have a statistically-significant association with fetal aneuploidy when seen in isolation. This document deals with the adjustment in risk for fetal aneuploidy based on the presence or absence of second trimester ultrasound markers; however, this risk adjustment has not been validated in a population with a lower prevalence for fetal aneuploidy owing to first trimester prenatal screening and diagnosis. As early screening (nuchal translucency, early maternal serum testing) and diagnosis (chorionic villus sampling) become established, the significance of second trimester markers will decrease and require readjustment.

In summary, the screening ultrasound at 16 to 20 weeks should evaluate 8 markers, of which 5 (thickened nuchal fold, echogenic bowel, mild ventriculomegaly, echogenic intracardiac focus, and choroid plexus cyst) are associated with an increased risk of fetal aneuploidy as well as nonchromosomal problems, while 3 (single umbilical artery, pylectasis, and enlarged cisterna magna) are only associated with an increased risk of nonchromosomal problems when seen in isolation.

References

EFW Services
Ultrasound
Maternal Fetal Medicine
Breast Imaging
Bone Densitometry
MRI
Nuclear Medicine
Pain Management

EFW Locations

Advanced Spinal Care Centre 2000 Veteran’s Place NW, Calgary
Advanced Medical Imaging Centre 2000 Veteran’s Place NW, Calgary
Airdrie Clinic #204, 836 1 Avenue NW, Airdrie
Beddington MFM and Clinic #200, 8120 Beddington Blvd NE, Calgary
Foothills Professional Building 1620 29th Street NW, Calgary
Gulf Canada Square #300, 401 9th Avenue SW, Calgary
Maternal Fetal Medicine Centre #100, 3280 Hospital Drive NW, Calgary
Nuclear Cardiology #210, 1016 68th Avenue SW, Calgary
Prostate Cancer Institute 6500 – 7007 14th Street SW, Calgary (at RGH)
Quarry Park MFM #130, 109 Quarry Park Blvd SE, Calgary
Southport Atrium Clinic A8, 10333 Southport Road SW, Calgary

EFW Contacts

**Physicians-Only Phone** (403) 210-9980
or (403) 944-1969

Central Booking (403) 541-1200
MFM Booking (403) 289-9269
Advanced Spinal Care & Pain Management (403) 244-3700
MRI Booking (403) 210-9090
中文预约电话 (403) 295-1880

Updated February 2013