Recommendations from international clinical guidelines for routine antenatal infection screening: does evidence matter?

Brigitte Piso, Inanna Reinsperger and Roman Winkler

Department of Public Health and Health Services Research, Ludwig Boltzmann Institute for Health Technology Assessment, Vienna, Austria

ABSTRACT

Aim: Maternal infections in pregnancy may cause severe child morbidity. In this article, we aim to summarise recommendations from international evidence-based clinical guidelines for infection screening in pregnancy.

Methods: We conducted a systematic search for evidence-based guidelines in two databases (Guidelines International Network and National Guideline Clearinghouse) and a hand search on websites of international institutions and societies that develop evidence-based guidelines. We considered guidelines published from the 1st of January 2007 onwards and developed in Western, industrialised countries. The guidelines in our analysis had to be based on a systematic literature search in at least two databases and needed to provide recommendations explicitly linked to the evidence. We included five general antenatal clinical practice guidelines from Australia, UK and the United States and 20 disease-specific guidelines published by Canadian, German, UK and US guideline development groups.

Results: The comparison of evidence-based guidelines from nine different guideline development groups for 17 antenatal infection screenings showed variations in directions (pro-screening or contra-screening) as well as in grades of recommendations. Uniformly, all-pro-universal or all-contra-universal screening recommendations were identified for 10 of 17 diseases. Contradictions were primarily observed for group B streptococcus, chlamydia trachomatis, genital herpes and gonorrhoea infection screening. Whereas certainty of recommendations was high for all-pro-screening recommendations, it decreased in all-contra-screening recommendations and even displayed conflicting results for contradictory recommendations.

Conclusion: The variety of grades of recommendations hamper across-guideline comparison. Nevertheless, the article highlights agreed screening areas based on the best available evidence as well as areas of still existing uncertainty. Local health policy decisions on whether to include or refrain from including screening measures in preventive care programmes can be facilitated by the comparison of recommendations from international evidence-based guidelines. Beyond the availability of evidence each country’s health policy makers will have to make a judgement on the value of the test for a population-wide screening.

Key words: communicable diseases, evidence-based practice, practice guideline, pregnancy, prenatal diagnosis

Background

Antenatal preventive care programmes show a wide variation of included screening measures across countries. Most of western (European) countries offer routine antenatal care to all pregnant women. The Regional Office for Europe of the WHO concluded back in 2003 that excessive, unneeded and unproven interventions are often provided to women with normal pregnancies. Against the backdrop of efforts towards the EU Directive on cross-border healthcare, Bernloehr et al. conducted a survey among European Ministries of Health and equivalent bodies, including European Societies of Obstetricians and Midwives in 2004. The
intention of this review of national guidelines was to find out whether a common European minimum guideline would be feasible and which essential parts of antenatal care this guideline might contain. Twenty member states provided information about their national antenatal care guidelines. The guideline overview showed that only three out of 37 screening tests (blood pressure, blood group and rhesus factor) were recommended in all 20 countries. The number of the provided laboratory tests ranged from 8 to 21.

The significance of maternal infections in pregnancy is evident. Maternal infections may cause severe child morbidity. Vertically transmitted infections [e.g. cytomegalovirus (CMV), rubella, toxoplasmosis, syphilis or varicella] may lead to malformations, neurodevelopmental delay and long-term childhood consequences. Maternal genital infections may increase the risk for miscarriage or preterm birth (e.g. vaginal bacteriosis) or cause neonatal infection by intrapartal transmission [e.g. group B streptococcus (GBS) infection or genital herpes]. Hence, the main rationale for infection screening during pregnancy is to reduce fetal or neonatal infections by early treatment of the infected pregnant woman. Another justification might be to enhance targeted preventative measures during the current pregnancy (e.g. caesarean section) or with regard to subsequent pregnancies (e.g. postpartum varicella or rubella immunisation).

Recommendations from clinical practice guidelines (CPGs) represent an initial point for discussions whether to introduce an intervention into a national preventive care programme or not. CPGs are defined as ‘systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical circumstances’. Systematic reviews of the effects of one intervention compared to another give information about the quality of evidence (e.g. high or low quality), whereas practice guidelines additionally involve decisive information regarding the strengths of screening recommendations. Most likely, convincing (high quality) evidence for the efficacy of a measure will lead to a recommendation to provide this measure (for a defined population wherein efficacy has been shown). Usually, additional considerations, for example patient-relevant outcomes, (national) burden of disease, the relation between benefits and harms and health economic factors (e.g. cost-effectiveness or the availability of technical/professional resources) influence recommendations drawn from evidence.

In Austria, there is an ongoing discussion whether the current mother-child preventive care programme reflects the changing needs of the target populations.

In the absence of a national screening committee or regulatory body, the ‘Ludwig Boltzmann Institute for Health Technology Assessment’ developed several decision support documents for the Austrian Ministry of Health for a potential re-orientation of the mother-child preventive care programme. Within this context, we conducted an overview of recommendations from evidence-based guidelines for screening measures during pregnancy and early childhood.

Aims
In this article, we aim to systematically summarise recommendations for maternal infection screening in pregnancy from international evidence-based antenatal guidelines (which cover prenatal care in general) as well as disease-specific clinical guidelines.

Methods
Search strategy
In June 2012, we conducted a systematic search for evidence-based guidelines in the databases of the Guidelines International Network (GIN) and the National Guideline Clearinghouse (NGC). To identify as many relevant guidelines as possible, we chose a broad search strategy: in the database of the GIN, we used ‘screening’ as the search term; in the database of the NGC, we also chose ‘screening’ as guideline category in the advanced search. Additionally, in July 2012 and 2013, we searched for relevant guidelines on websites of selected international institutions and societies that develop evidence-based guidelines or provide policy recommendations (Table 1). All references were exported to a software tool for managing bibliographies (Endnote X5, Thomson Reuters).

Inclusion criteria
We considered guidelines published from 1 January 2007 onwards and prior guidelines if guideline development groups (GDGs) reaffirmed the guideline validity (Table 2). In this context, we followed the respective inclusion criteria of the NGC. Accordingly, the NGC database exclusively involves guidelines that were developed, reviewed or revised over the past 5 years. Furthermore, we solely included guidelines from western, industrialised countries. The guidelines also needed to inform about the systematic literature search in at least two databases (to minimise retrieval bias). The grades of screening recommendations had to be explicitly linked to the underlying evidence (given rating schemes for the level of evidence and grades of recommendation). We refrained from using a specific checklist for quality assessment.
Two researchers assessed the references within a three-step selection procedure. First, we excluded references on the basis of guidelines’ titles due to language reasons (i.e., the guideline was not in English or German), to the geographical scope (e.g., developing country) or to thematic issues (e.g., oncology). Second, we screened summaries of the remaining references and excluded those which did not meet the inclusion criteria or which made recommendations exclusively for at-risk or ill women (see Table 2). Finally, the full-text guidelines of the remaining references were reassessed with regard to the above-mentioned inclusion criteria. Differences were solved by discussion and consensus or the involvement of a third researcher.

Data extraction
One researcher extracted the ‘direction’ of the screening recommendation and the respective grade of recommendation (in parentheses) for each screening measure from each of the included guidelines into a predefined extraction table. A second researcher assessed the extracted information for completeness and correctness. Regarding the ‘direction’ of the recommendation, we assigned a check mark ‘✓’ to a pro-screening, an ‘X’ to a contra-screening recommendation and a tilde ‘~’ if no recommendation was made (e.g., because of lack of or conflicting evidence). If screening was only recommended for a defined ‘at-risk’ group of pregnant women, we put the check mark in square brackets ‘[✓]’ (Table 3).

To compare the recommendations’ rating schemes, we extracted the notations and definitions of grades of recommendation from guideline development manuals or, if not available, from single guideline publications. We provide this additional information in a separate table (Table 4).

Results
Description of included guidelines
The database search for screening guidelines yielded a total of 655 references. After adding the references...
Table 3. Recommendations from evidence-based guidelines for infection screening in pregnancy

<table>
<thead>
<tr>
<th>Disease/test for</th>
<th>Pro-universal screening recommendations</th>
<th>Contra-universal screening recommendations (except for at-risk groups)</th>
<th>Contradictory universal screening recommendations</th>
</tr>
</thead>
</table>

**Pro-universal screening recommendations**

- **Hepatitis B virus**
  - AUS: (A)
  - UK: (A)
  - US: (NGR)
  - CAN: (A)
  - NG: (A)
  - Directions: 7

- **Asymptomatic bacteriuria**
  - AUS: (A)
  - UK: (NGR)
  - US: (A)
  - CAN: (A)
  - NG: (A)
  - Directions: 6

- **HIV**
  - AUS: (B)
  - UK: (A)
  - US: (NGR)
  - CAN: (A)
  - NG: (A)
  - Directions: 6

- **Syphilis**
  - AUS: (B)
  - UK: (A)
  - US: (NGR)
  - CAN: (B)
  - NG: (A)
  - Directions: 5

- **Rubella susceptibility**
  - AUS: (B)
  - UK: (B)
  - US: (NGR)
  - CAN: (B)
  - NG: (A)
  - Directions: 4

- **Varicella susceptibility**
  - AUS: NR
  - UK: NR
  - US: (NGR)
  - CAN: (NGR)
  - NG: (NG)
  - Directions: 1

**Contra-universal screening recommendations**

- **Bacterial vaginosis**
  - AUS: (B)
  - UK: (A)
  - US: (NGR)
  - CAN: (A)
  - NG: (NG)
  - Directions: 6

- **Cytomegalovirus**
  - AUS: (B)
  - UK: (A)
  - US: (NGR)
  - CAN: (A)
  - NG: (NG)
  - Directions: 5

- **Toxoplasmosis**
  - AUS: (B)
  - UK: (B)
  - US: (NGR)
  - CAN: (B)
  - NG: (A)
  - Directions: 3

- **Hepatitis C virus**
  - AUS: (C)
  - UK: (C)
  - US: (NGR)
  - CAN: (A)
  - NG: (NG)
  - Directions: 2

- **Parvovirus**
  - AUS: NR
  - UK: (C)
  - US: (NGR)
  - CAN: (A)
  - NG: (NG)
  - Directions: 1

- **Trichomoniasis**
  - AUS: (B)
  - UK: (A)
  - US: (NGR)
  - CAN: (A)
  - NG: (NG)
  - Directions: 5

**Contradictory universal screening recommendations**

- **Group B streptococcus**
  - AUS: (C) or RF-based prevention
  - UK: (C)
  - US: (NGR)
  - CAN: (A)
  - NG: (NG)
  - Directions: 2

**COMMENTARY**

International Journal of Evidence-Based Healthcare © 2014 University of Adelaide, Joanna Briggs Institute. Unauthorized reproduction of this article is prohibited.
### Table 3. (Continued)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Genital herpes</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>Gonorrhoea</td>
<td>X (cons)</td>
<td>at risk/ high prex</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

**Notes:**
- NR: no recommendation available; NG: no guideline available; 'direction': recommendation for or against a screening measure. ACOG, American College of Obstetricians and Gynecologists; AHMAC, Australian Health Ministers’ Advisory Council; AUS, Australia; AWMF, Association of the Scientific Medical Societies; CAN, Canada; GER, Germany; ICSI, Institute for Clinical Systems Improvement; NICE, National Institute for Health and Clinical Excellence; SOGC, Society of Obstetricians and Gynaecologists of Canada; RCOG, Royal College of Obstetricians and Gynaecologists; RF, risk factor; UK, United Kingdom; US, United States of America; USPSTF, United States Preventive Services Task Force; VA/DoD, Department of Veteran Affairs/Department of Defense.

- X: universal screening recommended for pregnant women.
- ✓: screening only recommended for a defined ‘at-risk’ group of pregnant women.
- ✓✓✓: universal screening not recommended for pregnant women.

- In parentheses: (grade of recommendation); (NGR) no grade of recommendation reported.
- (cons.) recommendation based on consensus.

- (a): number of recommendations identified per infection screening.
- (b): Recommendation extracted from consultation draft of guideline.
- (c): Available guideline included, despite literature search in single database: systematic search supported by comprehensive hand search for grey literature.
- (d): Guideline/policy is currently being reviewed.
- (e): Available guideline excluded: recommendation for preconceptional testing.
- (g): Available guideline excluded: literature search in single database.
### Table 4. Comparison of different grading schemes

<table>
<thead>
<tr>
<th>Grade of recommendation ‘A’</th>
<th>Recommendations are based on good and consistent scientific evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACOG12-34</td>
<td>Recommendations are based on good and consistent scientific evidence</td>
</tr>
<tr>
<td>AHMAC10,11</td>
<td>Body of evidence can be trusted to guide practice</td>
</tr>
<tr>
<td>AWMF20–22</td>
<td>Body of evidence can be trusted to guide practice</td>
</tr>
<tr>
<td>ICSI13</td>
<td>Previous class A: randomised controlled trial; now: high, if no limitation</td>
</tr>
<tr>
<td>NICE12</td>
<td>Directly based on level I evidence</td>
</tr>
<tr>
<td>RCOG35</td>
<td>At least one meta-analysis, systematic review or randomised controlled trial rated as 1++, and directly applicable to the target population; or a body of evidence consisting principally of studies rated as 1+, directly applicable to the target population, and demonstrating overall (showing considerable) consistency of results (with each other)</td>
</tr>
<tr>
<td>SCOG15–19</td>
<td>There is fair evidence to recommend the clinical preventive action</td>
</tr>
<tr>
<td>USPSTF36</td>
<td>The USPSTF recommends X service for Y population. Grade A indicates that the certainty of evidence is high that the magnitude of net benefits is substantial</td>
</tr>
<tr>
<td>VA/DoD14</td>
<td>A strong recommendation that the clinicians provide the intervention to eligible patients. Good evidence was found that the intervention improves important health outcomes and concludes that benefits substantially outweigh harm</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Grade of recommendation ‘B’</th>
<th>Recommendations are based on limited or inconsistent scientific evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACOG12-34</td>
<td>Recommendations are based on limited or inconsistent scientific evidence</td>
</tr>
<tr>
<td>AHMAC10,11</td>
<td>Body of evidence can be trusted to guide practice in most situations</td>
</tr>
<tr>
<td>AWMF20–22</td>
<td>Body of evidence can be trusted to guide practice in most situations</td>
</tr>
<tr>
<td>ICSI13</td>
<td>Previous class B: (observational) Cohort study; now: low</td>
</tr>
<tr>
<td>NICE12</td>
<td>Directly based on level II evidence or extrapolated recommendation from level I evidence</td>
</tr>
<tr>
<td>RCOG35</td>
<td>A body of evidence including studies rated as 2++ directly applicable to the target population and demonstrating overall consistency of results (considerable consistency with each other) OR extrapolated evidence from studies rated as 1++ or 1+</td>
</tr>
<tr>
<td>SCOG15–19</td>
<td>There is fair evidence to recommend the clinical preventive action</td>
</tr>
<tr>
<td>USPSTF36</td>
<td>The USPSTF recommends X service for Y population. Grade B indicates that the certainty of evidence is moderate that the magnitude of net benefits is either moderate or substantial, or that the certainty of evidence is high that the magnitude of net benefits is moderate</td>
</tr>
<tr>
<td>VA/DoD14</td>
<td>A recommendation that clinicians provide (the service) to eligible patients. At least fair evidence was found that the intervention improves health outcomes and concludes that benefits outweigh harm</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Grade of recommendation ‘C’</th>
<th>Recommendations are based primarily on consensus and expert opinion</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACOG12-34</td>
<td>Recommendations are based primarily on consensus and expert opinion</td>
</tr>
<tr>
<td>AHMAC10,11</td>
<td>Body of evidence provides some support for recommendation(s), but care should be taken in its application</td>
</tr>
<tr>
<td>AWMF20–22</td>
<td>Body of evidence provides some support for recommendation(s), but care should be taken in its application</td>
</tr>
<tr>
<td>ICSI13</td>
<td>Previous class C: (observational) nonrandomised trial with concurrent or historical controls. Case–control study, population-based descriptive study, study of sensitivity and specificity of a diagnostic test; now: low</td>
</tr>
<tr>
<td>NICE12</td>
<td>Directly based on level III evidence or extrapolated recommendation from level I or II evidence</td>
</tr>
<tr>
<td>RCOG35</td>
<td>A body of evidence including studies rated as 2++ directly applicable to the target population and demonstrating overall consistency of results (showing considerable consistency with each other) OR extrapolated evidence from studies rated as 2++</td>
</tr>
<tr>
<td>SCOG15–19</td>
<td>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>USPSTF36</td>
<td>The USPSTF recommends against routinely (providing) X service for Y population. There may be considerations that support (providing) the service in an individual patient. Grade C indicates that the certainty of the evidence is either high or moderate that the magnitude of net benefits is small</td>
</tr>
<tr>
<td>VA/DoD14</td>
<td>No recommendation for or against the routine provision of the intervention is made. At least fair evidence was found that the intervention can improve health outcomes, but concludes that the balance of benefits and harms is too close to justify a general recommendation</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Grade of recommendation ‘D’</th>
<th>[No statement possible (in general based on level III-V evidence)] or (based on level IV evidence); NA in20</th>
</tr>
</thead>
<tbody>
<tr>
<td>AWMF20–22</td>
<td>[No statement possible (in general based on level III-V evidence)] or (based on level IV evidence); NA in20</td>
</tr>
<tr>
<td>ICSI13</td>
<td>Previous class D: (observational) cross-sectional study, case series, case report; now: low</td>
</tr>
</tbody>
</table>

©2014 University of Adelaide, Joanna Briggs Institute. Unauthorized reproduction of this article is prohibited.
identified by supplementary search and subsequent removal of duplicates. 570 references were screened according to the three-step selection process described above. We finally included five antenatal CPGs (which cover prenatal care in general) from Australia,\cite{10,11} Great Britain,\cite{12} and the United States of America\cite{13,14} and 20 disease-specific guidelines published by Canadian,\cite{15-19} German,\cite{20-22} British,\cite{23-25} and American\cite{26-34} CPGs (Table 3). A total of 75 screening recommendations (53 from general antenatal and 22 from disease-specific guidelines) targeting 17 antenatal infections could be identified.

For each infection screening, we identified recommendations from one to seven GDGs. Most recommendations were made for hepatitis B screening (seven institutions),\cite{10,12-14,20,26,32} followed by screening for asymptomatic bacteriuria, HIV, bacterial vaginosis and CMV. Only one to two GDGs targeted recommendations for parvovirus,\cite{13,14} tuberculosis,\cite{13,14} varicella susceptibility\cite{14} and trichomoniasis\cite{11} screening. For seven diseases, we identified contra-recommendations for universal screening of all pregnant women (or women at low risk), but additional differing pro-recommendations to screen ‘at-risk’ populations.
Uniform (all-pro or all-contra) recommendations

Uniform recommendations were found for the majority of infection screenings. Unrestricted, ‘all-pro’-universal screening recommendations were found for hepatitis B,10,12–14,20,26,32 asymptomatic bacteriuria,10,12–14,21,27 HIV,10,12–14,23,28 syphilis10,12–14,29 and rubella susceptibility.10,12,14,16 For varicella susceptibility screening, only one single pro-screening recommendation could be identified.14 ‘All-contra’-universal screening recommendations could be identified for bacterial vaginosis,10,12–15,30 toxoplasmosis,10,12–14,18,33 hepatitis C,10,12,22,32 and parovirus.13,14 However, except for parovirus, additional pro-screening recommendations for defined ‘at-risk’ populations were formulated (bacterial vaginosis13,15,30, toxoplasmosis,18 hepatitis C).13,22,32 This also applies to the contra-screening tendency for CMV infection, which was not recommended by four GDGs11,12,17,33 (except for at-risk groups);11,13,17 one institution concluded that evidence was insufficient to recommend for or against routine screening.14 For trichomoniasis, we identified a contra-universal screening recommendation from a single GDG, but a recommendation to offer (diagnostic) testing to women who have symptoms.10 For tuberculosis, no universal screening recommendation was formulated; only ‘at-risk’ screening recommendations exist.13,14

Contradictory recommendations

Contradictory recommendations were found for GBS, chlamydia trachomatis, genital herpes (herpes simplex virus, HSV) and gonorrhoea infection screening. GBS screening obtained universal pro-screening recommendations from Australia13 and the United States,13,14 whereas two British GDGs12,24 militated against screening. Both chlamydial and HSV infection screening received three contra-universal and one pro-universal screening recommendation. For chlamydial infection, only one American GDG recommends a universal screening,14 whereas all other GDGs (from America,10,28 Great Britain9 and Australia7) stated contra-screening recommendations only for women younger than 25 years or at risk of infection. For genital herpes, the single pro-recommendation from Canada19 suggests a universal screening by history, but a laboratory screening only for women with a partner with HSV infection. Finally, for gonorrhoea screening, one American universal pro-screening recommendation14 is accompanied by an ‘at-risk groups only’ screening recommendation from a different American GDG10 and a contra-universal screening recommendation from Australia (including an ‘at-risk groups’ screening recommendation).11

Recommendation grades

GDGs did not use uniform grading systems (see Table 4). Nevertheless, grades of recommendations were nearly homogenous for ‘all-pro’-screening recommendations. In this category, most ‘A’ grades were observed. In general, ‘A’ grades reflect highest possible certainty to recommend the service. For hepatitis B and asymptomatic bacteriuria screening, GDGs solely assigned ‘A’ grades to the recommendation.

This high level of certainty and at least moderate uniformity of grades could not be observed for ‘all-contra’-screening recommendations. The recommendations against universal bacterial vaginosis screening are based on level I evidence (A),12 a body of evidence that can be trusted (B),7 at least moderate-to-high certainty of no net benefit (D)10 or fair evidence for ineffectiveness (D).14 The contra-hepatitis C-screening recommendation refers to limited or inconsistent scientific evidence (B),12 is just based on level III evidence (C)12 or a body of evidence that provides some support for recommendation(s), but care should be taken in its application (C).10 A disagreement in grade assignment can also be seen in toxoplasmosis screening recommendations: whereas one American GDG states that their recommendation is primarily based on consensus and expert opinion (C),13 a second one reports at least fair evidence of ineffectiveness (D)14, like a British GDG, which stated that the recommendation was directly based on level II evidence (B).12

The assigned grades of contradictory recommendations seem to be even more conflicting. GBS screening, for example, should be provided because at least fair evidence was found for health improvement (B) according to an American GDG.14 By contrast, British GDGs found level III or IV evidence to militate against a screening recommendation (with grades C12 or D24).

Discussion

The comparison of evidence-based guidelines from nine different institutions for 17 antenatal infection screenings showed variations in directions and grades of recommendations. Uniformly, all-pro-universal or all-contra-universal screening recommendations were identified for the majority of diseases (10 of 17). Contradictions were primarily observed for GBS, chlamydia trachomatis, HSV and gonorrhoea infection screening. Whereas the certainty of recommendations was high for all-pro-screening recommendations, it decreased in all contra-screening recommendations and even displayed conflicting results for contradictory recommendations.
Uniform pro-recommendations correspond largely to European screening recommendations surveyed in 2004\textsuperscript{3} and current UK National Screening Committee’s (UKNSC) policies.\textsuperscript{37} For example, all five uniform recommendations for the universal screening of pregnant women (for asymptomatic bacteriuria, hepatitis B, HIV, syphilis, rubella susceptibility) are found among recommended screening tests in 60–95% of the European countries in 2004.\textsuperscript{3} Other than screening for asymptomatic bacteriuria, for which a systematic population screening is not recommended by the UKNSC\textsuperscript{38} (though testing for asymptomatic bacteriuria within antenatal clinics should continue as part of good clinical practice as recommended in the National Institute for Health and Clinical Excellence (NICE) antenatal care guideline\textsuperscript{12}), the remaining four screening measures are also supported by the UKNSC on the population level.\textsuperscript{39–42} On the contrary, 35–50% of European countries recommended the hepatitis C virus or toxoplasmosis screening (in 2004),\textsuperscript{3} though – except for at-risk group pro-screening recommendations – only uniform universal contrascreening recommendations were found in the guideline overview. Available UKNSC ‘no population-based screening’ policy recommendations are in line with these uniform universal contra-screening recommendations.\textsuperscript{33,44}

Owing to the variety of grades of recommendations, across-guideline comparison is restricted. Grade ‘A’ in most grading schemes the highest grade of recommendation (see Table 4), is allotted to a recommendation if the body of evidence can be trusted.\textsuperscript{10,11} There is high certainty that the net benefit is substantial\textsuperscript{36} if it outweighs harm\textsuperscript{14} or there is (consistent) level I evidence (which might be a systematic review/meta-analysis of randomised controlled trials, or at least one randomised controlled trial).\textsuperscript{12,13,20–22,35} The handling of weak, insufficient or conflicting evidence seems to be nonuniform. If one GDG makes no recommendation for or against the routine provision of the intervention, it might assign a ‘C’ to the recommendation if at least fair evidence was found that the intervention can improve health outcomes, but concludes that the balance of benefits and harms is too close to justify a general recommendation or an ‘I’ if evidence that the intervention is effective is lacking, or of poor quality, or conflicting, and the balance of benefits and harms cannot be determined.\textsuperscript{14} Other GDGs also use special ‘I’\textsuperscript{36} or ‘L’\textsuperscript{15–19} notations if no recommendation is made because of lack of evidence. On the contrary, recommendations might be based primarily on consensus and expert opinion (‘C’,\textsuperscript{32–34} previous ‘class R or X’\textsuperscript{10}), the view of the GDG (‘Good practice point’ or ‘practice point’),\textsuperscript{10–12,35} or have been formulated in the absence of quality evidence (‘consensus-based recommendation’).\textsuperscript{10,11}

Limitations in our guideline comparison are as follows:

1. Potential differences in suggested screening points in time (e.g. weeks of gestation) and screening tests (e.g. methods of laboratory tests) were not analysed in depth for this guideline comparison. Cited pro-recommendation or contra-recommendations reflect general recommendations to screen for an infection irrespective of screening time or method.

2. At-risk recommendations apply for different, disease-specific risk groups of pregnant women. These range from women of a defined age group (chlamydial infection), child and healthcare workers in contact with young children (CMV), women at risk for preterm birth (bacterial vaginososis), women with a history of intravenous drug use or a history of blood transfusion prior to 1992 (hepatitis C), women who are immunosuppressed or HIV-positive (toxoplasmosis) or women with drug use, HIV, living in poverty or new immigrants from tuberculosis endemic areas (tuberculosis). Decisions on whether or not to screen pregnant women, therefore, depend on a multilevel approach. Precedent individual anamnestic risk assessment is, therefore, crucial to identify women at increased risk.

3. Because of the lack of a recent international compilation of nationally recommended or even reimbursed screening measures, no meaningful comparison of current guidelines and practice can be made. Instead, different sources, online available UKNSC policy recommendations\textsuperscript{27} and data from the European survey of 2004\textsuperscript{4} have been used as approximation.

Conclusion
Within the formulation of recommendations, evidence matters in varying degrees. At least under the assumption that the body of evidence did not change to a substantial extent within 5 years, its appraisal led to sometimes even contradictory decisions, especially if the level of evidence was low. Although all GDGs use defined schemes for grading the evidence levels and recommendations, factors that influence downgrading or upgrading of underlying evidence remain ambiguous in some aspects. Given the variety of grading systems used by different organisations, the GRADE working group suggested a uniform scheme to rate the quality...
of evidence and the strength of recommendations back in 2008. Some organisations already started to apply or at least adapt the suggested scheme. NICE has introduced the GRADE scheme for appraising evidence across its clinical guideline programme recently. First experiences show that the usage of GRADE necessitates a substantial shift from previous methods of evidence evaluation. GRADE facilitated the separation between judgments about the confidence in estimates of an effect from judgments about the strength of the corresponding recommendation. Nevertheless, new conceptual and practical problems arose, which will require further research (e.g. the application of GRADE to other than randomised trials).

A standardisation of grading schemes will be of low interest at local level, but might enhance joint efforts to provide the most appropriate and effective healthcare based on the most up-to-date and best-available evidence across countries.

Implications for research
The comparison of international evidence-based guidelines highlights agreed screening areas based on the best available evidence as well as infection screenings for which contradictory recommendations exist. In these areas, high-quality evidence is still lacking. Basically, screening targets a ‘healthy population’, because people undergoing screening at least believe that they are healthy with respect to the disease that is screened for. Taking into account that screening also involves the potential of harm (e.g. due to overdiagnosis, overtreatment, medicalisation), the relation between benefits and harms needs to be known when screening measures are implemented on a population level. Time trend and case-control studies might be useful in monitoring the effect of newly implemented screening measures or to compare different policies, but are at risk of various sources of bias and confounding factors. Despite obstacles, which complicate the completion of randomised controlled screening trials (e.g. recruitment of a healthy target population, large sample sizes if the assumed effect is small and subsequently high costs), well designed randomised controlled trials remain – according to Raffle and Gray – the only source of evidence about the consequences of screening.

Implications for practice
For clinical practice, debates about levels of evidence and grades of recommendation may play a secondary role in everyday decision-making.

Local health policy decisions on whether to include or refrain from including screening measures in preventive care programmes can be facilitated by the comparison of recommendations from international evidence-based guidelines, especially if the evidence for effectiveness – a beneficial relation between benefits and harm – is strong. Beyond the availability of evidence (weak or strong), each country’s health policy makers will have to make a judgement on the value of the test for a population-wide screening. Checklists like the 22 ‘criteria for appraising the viability, effectiveness and appropriateness of a screening programme’ used by the UKNSC can assist this decision-making process. Characteristics to be analysed cover universal criteria like test performance measures (e.g. sensitivity and specificity), but also population and healthcare system-specific criteria. The latter include an appraisal of the disease as a major public health problem (e.g. national epidemiology rates), the acceptance of the screening test, the costs of the screening programme as well as adequate staffing and facilities for testing, diagnosis, treatment and programme management (the availability of appropriate referral and management pathways). If this appraisal is transparent and comprehensible, the role of underlying evidence will become more evident.

Acknowledgements
The authors would like to thank Tarquin Mittermayr, BA, for conducting the systematic guideline search.

The Ludwig Boltzmann Institute for Health Technology Assessment is funded by national research funds (The Ludwig Boltzmann Gesellschaft) and by partner institutions (e.g. the Federation of Austrian Social Insurance Institutions and the Austrian Ministry of Health). This in-depth analysis of infection screening recommendations was carried out as decision support for the Austrian Ministry of Health within a larger project aiming at the re-orientation of the Austrian parent-child preventive care programme.

There are no conflicts of interest.

References