1

CHAPTER 1

Evidence-based medicine in obstetrics and gynecology

Jeanne-Marie Guise
Division of Maternal-Fetal Medicine, Departments of Obstetrics and Gynecology, Medical Informatics and Clinical Epidemiology, Public Health and Preventive Medicine, and Emergency Medicine, Oregon Health and Science University, Portland, OR, USA

> … decisions about the care of individual patients should be based on the conscientious, explicit, and judicious use of the current best evidence on the effectiveness of clinical services.
>
> *IOM Knowing What Works in Health Care 2008 [1]*

While all clinicians want to use the best evidence to make health care decisions, with 37 reviews, 47 randomized control trials (RCTs), and two guidelines published every day, it is impossible for practicing clinicians to keep up with all the new evidence and decide whether it is sufficient to suggest that they should change their practice. This book provides a summary of evidence for the major clinical areas of practice within the specialty of Obstetrics and Gynecology (OB/GYN), and this chapter (i) provides an overview and context, discussing the history of evidence based medicine (EBM) in OB/GYN; (ii) describes the importance and conduct of a systematic evidence review, a hallmark of EBM and contemporary evidence-based decision-making; and (iii) provides additional EBM resources and references for interested readers.

History of obstetrics and evidence-based medicine

OB/GYN has played a long and important role in shaping what is known today as EBM, although it did not always embrace evidence. The beginnings of OB/GYNs relationship with EBM may have started in the 1800s when women went to Lying-in Hospitals to stay for days or months in preparation for and recovery from childbirth. Lying-in hospitals were often crowded, and rates of maternal and child death from childbed fever (puerperal sepsis) were high. Some women were said to prefer giving birth in the streets, pretending to have given birth *en route* to the hospital. Ignac Semmelweiss, perplexed by the lower rates of maternal mortality for women delivering outside the hospital said: "To me, it appeared logical that patients who experienced street births would become ill at least as frequently as those who delivered in the clinic … What protected those who delivered outside the clinic from these destructive unknown endemic influences?" [2]. He also observed that there were higher rates of maternal mortality from childbed fever in the First Division Hospital, which was staffed by physicians, compared with the Second which was staffed by midwives. Both units had trainees, performed examinations, and saw roughly similar populations. He realized that unlike the midwives, physicians all performed autopsies on women who died the night before prior to beginning their clinical duties on the maternity ward. In 1847, Semmelweiss figured out what might be occurring when a forensic medical professor, Jakob Kolletschka, died of sepsis after sustaining an accidental finger stick during an autopsy. He concluded that, "In Kolletschka, the specific causal factor was the cadaverous particles that were introduced into his vascular system. I was compelled to ask whether cadaverous particles had been introduced into the vascular systems of those patients whom I had seen die of this identical disease. I was forced to answer affirmatively" [2]. He required physicians wash their hands with chlorinated lime before examining patients. The mortality rate in District 1 fell from 11.4% prior to handwashing to 1.27% (rates were 2.7% and 1.33% in District 2). The medical community did not embrace this new evidence. Semmelweiss was ridiculed by physicians who were offended by the suggestion they were unclean, and his theory was rejected because it was contrary to the accepted belief that childbed fever was caused by miasmas or "bad air." In response, Semmelweiss could only figuratively shake his head: "One would believe that the clarity of things would have made the truth apparent to everyone and that

Evidence-Based Obstetrics and Gynecology, First Edition. Edited by Errol R. Norwitz, Carolyn M. Zelop, David A. Miller, and David L. Keefe.

they would have behaved accordingly. Experience teaches otherwise. Most medical lecture halls continue to resound with lectures on epidemic childbed fever and with discourses against my theories" [2].

Fast forward to the 1950s and 1960s and two stories demonstrate how difficult it is for new evidence to change clinical practice even when that evidence is strong – and how profound the consequences for this failure.

In the 1950s, diethylstilboestrol (DES) therapy was used to prevent miscarriage. Its use was established through uncontrolled studies. Even though randomized controlled trials were published in the mid-1950s that found no significant prevention offered by DES, its use had become so commonplace that it continued despite the evidence. It was not until 1971 that the food and drug administration (FDA) brought national attention to the harms of DES exposure (associated with vaginal clear cell carcinoma) and banned its use in pregnancy. Total exposure to DES for mothers and daughters has been estimated to exceed 10 million worldwide.

The story of antenatal corticosteroids is not only a major discovery in obstetrics but is also emblematic of the importance of EBM. In the 1960s, Graham "Mont" Liggins, an Australian obstetrician, had a sheep farmer neighbor and wondered why ewes delivered prematurely when worried by dogs [3]. Liggins suspected it may have something to do with the stress-response in the mother and the release of cortisol. He conducted an experiment where he administered corticosteroids to pregnant ewes and found they delivered prematurely. Unexpectedly, he also found that the lambs delivered by ewes that received corticosteroids survived in far greater numbers than he would have expected given the severe degree of their prematurity [4]. In the 1970s, Liggins and a pediatrician colleague, Ross Howie, conducted the first randomized trial in humans to test their theory that corticosteroids reduced the occurrence of respiratory distress syndrome (RDS). RDS and mortality rates were significantly reduced in the treated group (6.4%) as opposed to 18% in placebo treated mothers. Within a decade of this first RCT additional studies supported the conclusion that corticosteroids significantly reduced infant mortality for prematurely born children. However it was not until the mid-1990s that antenatal steroids became part of routine practice for women at risk of premature delivery (after a meta-analysis was published in 1989). The forest plot from a meta-analysis of antenatal corticosteroids represents this delay, demonstrates the potential power of systematic reviews and meta-analyses of a body of evidence, and has become the symbol for the Cochrane Collaboration, the most recognized source for evidence-based systematic reviews in medicine. It has been estimated that tens of thousands of babies would have been saved by earlier implementation of steroids.

It is perhaps not a surprise that Archie Cochrane, for whom the Cochrane Collaboration is named awarded the field of OB/GYN the first wooden spoon award for failing to evaluate the care they provide with RCTs and failing to apply results of RCTs in practice [5]. He went further stating that GO in Gynecology and Obstetrics should stand for "go ahead without evidence" [6].

What is evidence-based medicine?

EBM, refers to a process of turning clinical problems into questions and systematically locating, appraising, and synthesizing research findings as a basis for clinical decision-making. Gordon Guyatt [7] first used the term "EBM" in the 1980s to describe an approach to residency training at McMaster University School of Medicine where residents were taught how to identify, interpret, and use the literature in their clinical decision-making. At first he wanted to call it "Scientific Medicine" but reconsidered when others argued that the title would imply all other medicine was non-scientific [8]. Further refined by David Sackett, "EBM requires a bottom-up approach that integrates the best external evidence with individual clinical expertise and patient choice" [9].

The systematic review is a hallmark of EBM. Systematic reviews apply a scientific review strategy that limits bias by the systematic assembly, critical appraisal, and synthesis of all relevant studies on a specific topic. As shown in Figure 1.1, systematic reviews are at the top of the evidence hierarchy pyramid. Clinicians in pursuit of the best evidence, should first search for high-quality systematic reviews. Since systematic reviews are such an important part of EBM and are instrumental to clinical decision-making, this chapter provides a brief description of the systematic review process.

Systematic review processes

If, as is sometimes supposed, science consisted in nothing but the laborious accumulation of facts, it would soon come to a standstill, crushed, as it were, under its own weight... Two processes are thus at work side by side, the reception of new material and the digestion and assimilation of the old [10]

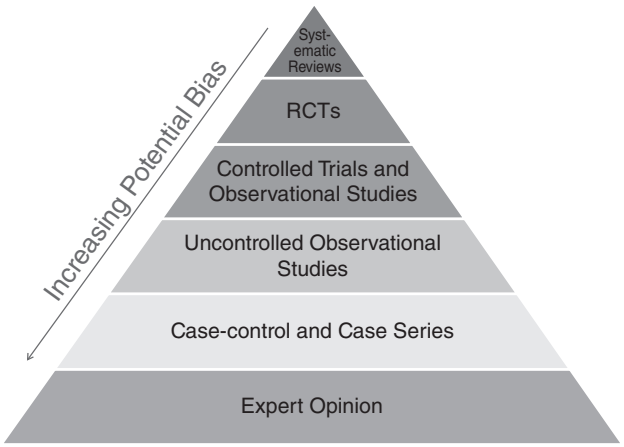

Figure 1.1 Systematic review processes.

Table 1.1 Steps for evidence-based obstetrics

1. Formulate a clear clinical question
2. Search the literature and identify relevant reviews and studies
3. Critically appraise individual studies and the overall body of evidence
4. Synthesize results given context and patient factors
5. Implement
6. Evaluate the application into clinical practice

A systematic review is a scientific review strategy that limits bias by the systematic assembly, critical appraisal, and synthesis of all relevant studies on a specific topic. Table 1.1 presents the six steps for Evidence-based Obstetrics. The first four of these are covered by, and critical to, systematic review. Therefore, busy clinicians can shortcut these steps if they are able to find a high-quality systematic review that answers their clinical question.

Each of these steps is covered briefly below.

Formulating the question

> A prudent question is one-half of wisdom [11]
>
> Sir Francis Bacon

Questions arise every day a clinician cares for patients: some they can answer easily, others they know where to find the answers quickly, and many require investigation. The ability to take an everyday dilemma and turn it into an answerable and searchable question is important not only for systematic reviews, but also for good clinical care. Questions often fall into specific categories: incidence/prevalence, causation/etiology, screening, diagnostic, therapeutic/treatment, prevention, outcomes (benefits and/or harms), prognostic, and they can be expressed as, "In patients with … how effective is … compared with … for the outcome[s] of … ". Formulating an answerable and relevant question is a critical foundational step to determining the relevant scope of a review; too big and the review may not be feasible, too narrow and the results may not be relevant. Systematic review questions are often formulated according to a PICOTS format, that is, **P**opulation, **I**ntervention, **C**omparator, **O**utcome, **T**iming, and **S**etting (Table 1.2).

Table 1.2 PICOTS

Population – Who does the review topic pertain to
Intervention – What is the intervention or treatment that is being evaluated?
Comparator – What is the intervention being compared with?
Outcome – What are the benefits and harms?
Timing – What is the timing of outcomes or follow-up?
Setting – What settings are relevant to this topic?

Population – Understanding the population of reviews and research studies is often one of the clearest ways clinicians can determine whether the scope of a review or study is pertinent to their clinical population. Factors often considered include age (e.g. child, teen, young adult, elderly), sex, medical conditions, pregnancy status, and social factors (education required, skill-level, access to care). A description of such factors helps clinicians understand whether the review will be applicable to their patient population.

Intervention – The intervention is often the main subject of reviews. Interventions can involve medical, surgical, health systems, social, or behavioral interventions and can have one or many components.

Comparator – The comparator group is often overlooked, yet is critical to understanding the relative effectiveness of an intervention. Comparators include no treatment, placebo, "standard of care," active alternative treatment. It is important to describe the underlying condition considered "standard of care" as what is considered standard might be an intervention in other settings.

Outcomes – Outcomes include health outcomes, intermediate outcomes, and harms.

Timing – Timing is increasingly recognized as an important consideration. Timing refers to the timing of the intervention or parts of the intervention and also may describe the time of patient eligibility, intervention, and follow-up for a target trial.

Setting – Setting or context factors such as organizational characteristics, financial setting (fee-for- service, capitated, uninsured; geographic and clinical settings (solo or group practice, public or private, for profit or non-profit, etc.) are often critical to interventional effectiveness and should be described in systematic reviews.

Often the S in PICOTS is used to refer to study design. While that use is not usually an element in the question, it can be helpful to consider the types of studies that are most likely to inform particular types of questions. Table 1.3 aligns common types of questions with study designs.

Descriptions of these PICOTS elements enables the reader of a systematic review to understand whether the question is

Table 1.3 Studies applicable to particular review questions

Question type	Study design
Incidence	Cohort
Prevalence	Cohort, cross-sectional
Treatment/therapy	Randomized controlled trial (RCT)
Screening	RCT, cohort
Diagnostic accuracy	RCT, case series
Prognosis	RCT, cohort
Harms	RCT, cohort, case–control, case report
Etiology	Cohort, case–control
Prevention	RCT, cohort

relevant to their clinical dilemma and setting. The questions also specify search terms and the inclusion and exclusion criteria for studies.

Searching the literature and identifying relevant studies

A comprehensive search and a systematic, unbiased approach to finding, selecting, and interpreting evidence are distinguishing features of systematic reviews. Searches of systematic reviews are meant to include all of the evidence and not just published articles. In general, bibliographic searches for systematic reviews in health care should always include MEDLINE® and the Cochrane Central Register of Controlled Trials. Additional databases that are often useful include Embase, CINAHL, Scopus, and PsychINFO. In addition to searching bibliographic databases, systematic reviews search reference lists of relevant reviews and articles and conduct searches for unpublished literature from registries, government or industry documents, Websites, and other sources. Once you have conducted a comprehensive search, the next critical ingredient of a systematic review is applying an unbiased approach to including and excluding articles. This process involves a priori decision-making about issues such as date range, study design, language, key subject matter issues etc. A PRISMA [12] or QUORUM [13] figure is often used to detail finding and selecting pertinent literature for a review.

Critically appraising studies and assessing the strength of a body of literature

Critically appraising the literature involves two major stages: (i) evaluating the risk of bias for individual studies based upon study design; and (ii) grading the overall strength of evidence for a body of literature. Problems with an individual study's design or conduct have the potential to introduce bias or inferential error, and raise questions about the validity of their findings. Numerous tools exist to evaluate the risk of bias for controlled trials [14–16] and observational studies [16–25]. In general risk of bias tools evaluate participant selection; outcome, exposure, and process measures; study processes such as blinding; and appropriate analytic methods including intent to treat and considerations for confounding. This stage of individual study evaluation is critically important. One element in assessing the strength of the body of literature, it can inform quantitative syntheses such as meta-analyses, and provide insights on how to strengthen future research studies in design and conduct. Because raters may vary in their interpretation, reviewers will usually pilot test the application of the tool prior to wide-scale use across studies.

Understanding the reliability of the overall body of evidence is critical for guideline groups, policymakers, and clinicians. Methods for evaluating the overall strength of evidence have evolved over the past several decades. Organizations such as the US Preventive Services Task Force (USPSTF) [19] the US Evidence-based Practice Centers (EPCs) Program [20], and the Oxford Center for Evidence-based Medicine [21] have all developed criteria. The USPSTF risk of bias/quality rating scale has been adapted for easy use by relative novices and is available at www.storc.org) In 2000, a collaboration of international experts formed the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) Working Group to establish common and transparent criteria to grade the literature. The group has grown tremendously over the years and experts in the field continue to refine the application of GRADE criteria by examining and debating their experiences and exemplars (www.gradeworkinggroup.org). According to GRADE, evidence from randomized controlled trials starts as high quality and that from observational studies starts as low quality based on the assumption that randomization controls for systematic bias in effect estimates. The body of evidence is evaluated using five main criteria: (i) risk of bias; (ii) inconsistency of results across studies; (iii) indirectness; (iv) imprecision; and (v) publication bias [22]. Risk of bias was discussed above. Consistency involves determining the degree to which studies were similar in direction and range of effect sizes. Directness involves assessing whether the evidence reflects a single direct link to the outcome or whether it involves several indirect links in a chain of evidence or surrogate outcomes. Precision has to do with the certainty of the effect which is often judged by the narrowness of the confidence interval. Publication bias is the last major GRADE criterion. It has long been recognized that studies with positive findings are more likely to be published. (Several factors can contribute to this, including journal bias toward positive results and author awareness of those journal preferences.) This alone can bias the overall body of literature. Published studies can show an intervention's effect while there could be a large body of unpublished evidence suggesting no effect. Because of this, GRADE recommends conducting an evaluation for publication bias. After considering GRADE elements, the entire body of literature for a given outcome is rated as high, moderate, low, or very low. Table 1.4 presents the summary grades and their meaning.

Knowing that guideline groups, policymakers, and clinicians have limited time, the GRADE working group also recommends use of a summary of evidence table to summarize: (i) key outcomes; (ii) effect sizes (magnitude and confidence interval); (iii) numbers of studies and participants; (iv) overall GRADE of evidence by outcome; and (v) important notes or comments. Ultimately, the GRADE approach provides a system for evaluating the strength of the literature as a whole and determining the strength of recommendation that can be made. For example a strong

Table 1.4 GRADING the quality of a body of literature [22]

High – Further research is very unlikely to change our confidence in the estimate of effect. (e.g. High confidence that the evidence reflects the true effect).
Moderate – Further research may change our confidence in the estimate of effect and may change the estimate. (e.g. Moderate confidence that the evidence reflects the true effect).
Low – Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate. (e.g. Low confidence that the evidence reflects the true effect).
Very low – any estimate of effect is very uncertain (e.g. very low confidence that the evidence reflects a true effect)

Table 1.5 List of evidence-based organizations and resources

Agency for Healthcare Research and Quality (AHRQ) – http://www.ahrq.gov
AHRQ Evidence-based Practice Centers Program (EPC) – http://www.ahrq.gov/research/findings/evidence-based-reports/index.html
Bandolier – http://www.medicine.ox.ac.uk/bandolier
Centre for Reviews and Dissemination (CRD) – www.york.ac.uk/crd
Cochrane Collaboration – http://www.cochrane.org
Cochrane Pregnancy and Childbirth – http://pregnancy.cochrane.org
Cochrane Gynecology and Fertility Group – http://cgf.cochrane.org
Cochrane Fertility Regulation Group – http://fertility-regulation.cochrane.org
Cochrane Gynecological Cancer Group – http://gnoc.cochrane.org
GRADE Working Group – http://www.gradeworkinggroup.org
JAMA Evidence – http://jamaevidence.mhmedical.com
James Lind – http://www.jameslindlibrary.org
National Institute for Health and Clinical Excellence (NICE) – www.nice.org.uk
Oxford Centre for Evidence-based Medicine – http://www.cebm.net
PRISMA – http://www.prisma-statement.org
US Preventive Services Task Force (USPSTF) – http://www.uspreventiveservicestaskforce.org

recommendation could be made when the effect size is large and overall evidence quality is high, meaning that it is unlikely to have occurred in the absence of a true effect of the intervention. However, a weak recommendation would be made for low or very low evidence where any effect could have occurred solely as a result of bias from confounding factors. The GRADE system or adaptations of the GRADE system are used by numerous guideline groups including since 2015 the International Consensus on cerebroplacental ratio (CPR) and endocervical curettage (ECC), Science with Treatment Recommendations provided by the International Liaison Committee for Resuscitation (ILCOR) which are used in this book [23]. Ultimately these processes and products are tools to promote transparency, understanding, and dialogue around the totality of evidence, our certainty in that evidence, and a rationale for practice.

Evidence-based resources

Table 1.5 provides the interested reader with additional resources to find evidence-based reviews and guidance and/or to learn more about evidence-based practices. Some of the major resources are discussed in some detail.

The Cochrane Collaboration

Realizing that it is a daunting if not impossible challenge for the individual practicing clinician to keep abreast and synthesize the medical literature, Sir Ian Chalmers, motivated by Archie Cochrane's wooden spoon challenge to obstetrics, developed a database of all existing and relevant randomized controlled clinical trials for interventions in OB/GYN and a repository of systematic reviews the Cochrane library. The Cochrane Collaboration (http://www.cochrane.org) is now one of the largest networks of global scientists, with more than 37 000 volunteers who synthesize the world's evidence and produce high-quality systematic reviews. The Collaboration is organized into review groups that are responsible for conducting and updating systematic reviews for specific topic areas. Several review groups are pertinent to OB/GYN including:

Pregnancy and Childbirth
The Cochrane Menstrual Disorders and Subfertility Group
The Cochrane Fertility Regulation Group
The Cochrane Gynecological Cancer Group

The Cochrane Library (http://www.cochranelibrary.com) has become one of the world's most recognized sources of high-quality systematic reviews in medicine. The origins and symbol of the Cochrane are connected to obstetrics, and as mentioned earlier, the very symbol for the Cochrane reflects the story of antenatal corticosteroid therapy.

The US preventive services task force and the US evidence-based practice centers program

The USPSTF (www.uspreventiveservicestaskforce.org) is an excellent resource for evidence and recommendations in primary care and prevention. The USPSTF was established in 1984 as an independent, volunteer panel of national experts in prevention and EBM who issue recommendations on clinical preventive services such as screenings, counseling services, and preventive medications. Topics relating to OB/GYN and women's health include cervical cancer screening; screening for bacterial vaginosis in pregnancy to prevent preterm birth; mammography; breastfeeding; screening for BRCA-related cancer, chlamydia, and gonorrhea, depression, genital herpes; counseling for

gynecologic cancers; immunizations, and many more. It is an excellent resource for primary care issues and is considered by the US government when making coverage decisions. All USPSTF recommendations are paired with systematic evidence reviews conducted by EPCs. In 1997, Agency for Healthcare Research and Quality (AHRQ) (then known as the Agency for Health Care Policy and Research) established the EPC program to develop evidence reports to inform health policy, guidelines, coverage decisions, patient decision-making, and clinical practice for clinical professional societies, insurers, employers, healthcare organizations, and policymakers. Examples of reports that are relevant to OB/GYN include comparative effectiveness of therapies to treat menopausal symptoms, antidepressant treatment of depression during pregnancy and postpartum, smoking cessation interventions in pregnancy and postpartum care, oral contraception use for the prevention of ovarian cancer, progestogens for the prevention of preterm birth, and nitrous oxide for the management of labor pain (a full list can be found at http://www.ahrq.gov/research/findings/evidence-based-reports/search.html?f[0]=field_evidence_based_reports%3A13971).

Rationale for this book

Clinicians have more access to evidence than ever before; this is both a cure and a curse. While the process of finding, appraising, and synthesizing evidence is possible for practitioners, studies suggest that the process is too time consuming for most [24, 25]. Inadequate time (74%), limited searching skills (41%), and limited access to evidence (43%) have been cited by physicians as barriers to implementing evidence-based care [25]. This book is written to provide a central resource for evidence in OB/GYN for the busy clinician. The chapters that follow provide an overview of the evidence across major clinical topics faced on a daily bases by Obstetricians and Gynecologists.

References

1 Eden, J., Levit, L., Berg, A. et al. (eds.) (2011). *Finding What Works in Health Care: Standards for Systematic Reviewers*. The National Academies Press.

2 Semmelweis, I. and Carter, K.C.T. (1983). *Etiology, Concept and Prophylaxis of Childbed Fever*. University of Wisconsin Press.

3 Norwitz, E. and Greeberg, J. (2010). Beyond antenatal corticosteroids: what did mont liggins teach us? *Reviews in Obstetrics and Gynecology* **3** (**3**): 79.

4 Gluckman, P. and Buklijas, T. (2013). Sir Graham Collingwood (Mont) Liggins. *Biographical Memoirs of Fellows of the Royal Society* **59**: 193–214.

5 Ohlsson, A. and Clark, K. (2002). Evidence-based perinatal and paediatric health care: the contribution of the Cochrane collaboration and tte Canadian Cochrane network and Centre. *Paediatrics & Child Health* **7**: 319–322.

6 The Cochrane Pregnancy and Childbirth Database. (1989). Effective Care in Pregnancy and Childbirth. http://journals.sagepub.com/doi/abs/10.1177/016327879501800406.

7 Guyatt, G. and Voelker, R. (2012). Everything you ever wanted to know about evidence-based medicine. *JAMA* **313** (**18**): 1783–1785.

8 Smith, R. and Rennie, D. (2014). Evidence-based medicine—an oral history. *JAMA* **311** (**4**): 365–367.

9 Sackett, D., Rosenberg, W., and Gray, J. (1996). Evidence based medicine: what it is and what it isn't. *BMJ* **312**: 71–72.

10 Raleigh, L. (1884). Address by the Rt. Hon. Lord Rayleigh. In: *Report of the Fifty-Fifty-Fourth Meeting of the British Association for the Advancement of Science*, 3–23. John Murray.

11 Bacon S.F. http://www.azquotes.com/quotes/topics/prudent.html

12 Moher, D., Liberati, A., Tetzlaff, J., and Altman, D. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of Internal Medicine* **151** (**4**): 264–269.

13 Clarke, M. (2000). The QUORUM statement. *Lancet* **355**: 756–757.

14 Collaboration TC (2011; 5.1.0). *Cochrane Handbook for Systematic Reviews of Interventions* (ed. H. JPT and S. Green). Wiley.

15 Higgins, J. and Altman, D. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* **343**: d5928.

16 Viswanathan M., Ansari M., Berkman N., et al. (2012) Assessing the Risk of Bias of Individual Studies in Systematic Reviews of Health Care Interventions Agency for Healthcare Research and Quality; Methods Guide for Comparative Effectiveness Reviews (AHRQ Publication No. 12-EHC047-EF).

17 Harris, R., Helfand, M., Woolf, S. et al. (2001). Methods work group, third US preventative services task force. Current methods of the US preventative services task force: a review of the process. *American Journal of Preventative Medicine* **20** (**3**): 21–35.

18 Wells G., Shea B., O'Connell D., et al. The Newcastle-Ottawa Scale (NOS) for Assessing the Quality of Nonrandomised Studies in Meta-Analyses https://www.ncbi.nlm.nih.gov/books/NBK99082/bin/appb-fm4.pdf.

19 Force USPST. Grade Definitions. 2015. https://www.uspreventiveservicestaskforce.org/Page/Name/grade-definitions.

20 Berkman, N.D.L.K., Ansari, M., McDonagh, M. et al. (2013). Grading the strength of a body of evidence when assessing health care interventions for the effective health care program of the Agency for Healthcare Research and Quality: an update. *Methods Guide for Comparative Effectiveness Reviews AHRQ* Publication No. 13(14)-EHC130-EF.

21 Howick, J., Chalmers, I., Glasziou, P. et al. (2011). The 2011 Oxford CEBM levels of evidence (introductory document). *Oxford Centre for Evidence-Based Medicine*.

22 Guyatt, G., Oxman, A., Vist, G. et al. (2008). Rating quality of evidence and strength of recommendations GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *British Medical Journal* **336**: 92.

23 Morrison, L., Chair, L., Gent, E. et al. (2015). Part 2: evidence evaluation and Management of Conflicts of interest 2015 American Heart Association guidelines update for cardiopulmonary resuscitation and emergency cardiovascular care. *Circulation* **132**: S368–S382.

24 Cumbers, B. (1999). Evidence-based practice data day. *Health Services Journal* **109** (**5650**): 30–31.

25 Craig, J., Irwig, L., and Stockler, M. (2001). Evidence-based tools: useful tools for decision making. *The Medical Journal of Australia* **174**: 248–253.