Evidence-based medicine in obstetrics: can levels B and C recommendations be elevated to level A recommendations?

For the Evidenced-Based Medicine Obstetric Group: Suneet P Chauhan MD*, Eugene Chang MD†, Brian Brost MD‡, Barbara Assel MD*, Jason Baxter MD§, James A Smith Jr. MD, Robert Grobman MD††, Vincenzo Berghella MD§, James A Scardo MD‡‡, Everett F Magann MD§§ and John C Morrison MD*****

*From Aurora Health Care, West Allis, WI; †Medical University of South Carolina, Charleston, SC; ‡Mayo Clinic, Rochester, MN; §Thomas Jefferson University, Philadelphia, MD; **Stanford University Medical Center, Stanford, CA; ††Northwestern University Medical Center, Chicago, IL; ‡‡Spartanburg Regional Medical Center, Spartanburg, SC; §§Portsmouth Naval Medical Center, Portsmouth, VA; ***University of Mississippi Medical Center, Jackson, MS, USA

Summary: In this study, 65% (132/195) of level B/C obstetric recommendations are amenable to randomized clinical trials (RCTs) and seven were identified as most needed. The purpose of the survey was to evaluate levels B and C recommendations in obstetric practice bulletins (PBs) regarding the feasibility of performing RCT to elevate each subject to level A evidence. Eleven geographically dispersed physicians with experience in research reviewed levels B and C recommendations for the ethical and logistical feasibility of performing an RCT. In the 35 obstetric PBs, 195 level B/C recommendations were reviewed. The majority considered 47 (24%) topics unethical for an RCT and thought 16 (11%) did not need an RCT, thus leaving 132 (67%) levels B and C recommendations available for an RCT. Two-thirds of levels B and C recommendations in obstetric PB are amenable to RCTs and potentially becoming level A evidence.

Keywords: ACOG obstetric bulletins, randomized clinical trials

INTRODUCTION

In every clinical practice, as well as academic teaching settings, it is important to objectively synthesize the literature and, whenever possible, make evidence-based recommendations for patient care scenarios. Practice bulletins (PBs) from the American College of Obstetricians and Gynecologist (ACOG) assist in this task by reviewing the literature on specific topics and then analysing the references according to the US Preventive Service Task Force criteria categorizing them as levels I, II or III evidence. The strength of the references leads to recommendations that are classified as A, B and C. Practice based on clinical consensus and expert opinions are level C recommendations, whereas level B recommendations are derived from limited or inconsistent scientific evidence while level A recommendations are based on consistent and scientific evidence.¹

Ideally, all ACOG recommendations should be level A; however, a previous review of 55 PBs published between 1998 and 2004 demonstrated that among 438 recommendations, 29% were level A, 33% level B and 38% level C.² One way to

ensure that most of the recommendations are of the highest quality is to perform Consolidated Standards of Reporting Trials (CONSORT)³ compliant randomized clinical trials (RCTs) on the topics that are levels B and C. However, due to ethical concerns or logistical difficulties, it may not always be possible to perform an RCT on a given subject. Moreover, it would be useful to prioritize levels B and C recommendations that urgently need an RCT.

The purpose of this survey was to evaluate all B- and C-level recommendations in obstetric PB and ascertain the feasibility and/or difficulties of conducting an RCT that would elevate the evidence to level A recommendation.

MATERIAL AND METHODS

Eleven obstetric specialists (maternal–fetal medicine [MFM] certified) geographically dispersed around the country who had extensive experience in RCTs formed the survey group. They were mailed all B- and C-level recommendations ($n = 195$) from 35 current (June 1998 to December 2005) obstetric PBs from ACOG. The reviewers were asked to answer five questions about each of them: (1) is it unethical to do an RCT on the topic?; (2) is an RCT needed regarding the recommendation?; (3) if an RCT is needed, how difficult would it be to accomplish (rate from 1 being easy and 5 hardest)?; (4) would it require (1), (2), (3) or the MFM network to conduct the RCT?; and (5) rate

Correspondence to: Suneet P Chauhan, MD, Aurora Health Care, 8901 W. Lincoln Avenue, PAC, West Allis, WI 53227, USA
Email: suneet.chauhan@aurora.org

the five topics that most urgently require an RCT. The review started in August 2006 and all the data were completed by May 2007.

To counter fatigue and its effect on the participants' answers, the recommendations were mailed in sets of 50. Once the participants returned their evaluations, they received the second set, a third set and finally a fourth set. Participants in the survey were not aware of the answers by other respondents and did not know if the recommendations they were evaluating were level B or C. After much discussion, we deliberately chose not to employ a universal framework for ethical reasons⁴ as it would be nearly impossible for our respondents to apply such criteria uniformly as they are subjective. Rather, we selected our participants based on their experience in research, clinical practice and teaching, as well as their interest in evidenced-based medicine. Our *a priori* assumption was that if the majority of the respondents found a proposed RCT to be unethical, it was unlikely to be approved through an Institutional Review Board. Determining the funding for the RCT was beyond the scope of the study because it would vary between topics, institutions and individuals conducting the study. A person not involved with the study entered all the data into an Excel spreadsheet. Odds ratio (OR), along with 95% confidence intervals (CI), was calculated and CI not crossing the integer 1 was considered significant.

RESULTS

Over a 90-month period, there were 36 PBs that focused on obstetric issues, but only one of them (ultrasonography) contained only level A recommendations and was excluded from the survey. Among the 35 bulletins eligible for survey there were 253 recommendations, of which 23% (58) were level A, 37% (94) were level B and 40% (101) were level C recommendations. Thus, 195 ACOG B/C recommendations formed the survey and each participant answered the 990 questions. The respondent's mean ($\pm$ standard deviation) age was 46.8 ± 9.0 years. The time lapsed between completion of the MFM fellowship and the survey was 10.9 ± 7.2 years. While two (18%) of the MFM surveyed were in private practice, 78% (8/11) were working at an academic centre and one (9%) in an Armed Forces Hospital, with a residency programme in obstetrics and gynaecology. All 11 had a consultative practice, were involved in teaching obstetrics and gynaecology residents, and seven (64%) were involved in the educating of MFM fellows. The mean number of publications by the respondents was 100 ± 148 .

The majority of participants considered it unethical to perform an RCT on 24% (47) of levels B and C recommendations (Figure 1). Excluding the 47 recommendations that thought it unethical to pursue, there were 148 potential topics available for a clinical trial. Of these 148 recommendations, the majority of MFMs (76–95%) thought 16 (11%) did not require an RCT, as the recommendations were currently accepted in clinical practice by the majority of obstetricians and thus the *de facto* constituted standard care (Table 1). Therefore, 132 (67% of all B- and C-level) recommendations were available for study (Figure 1). Of the 16, suggestions not needing RCT, 44% (7/16), were level B and nine were level C.

Of the 132 recommendations that were ethically justifiable and needed RCT, 52% (69) were level B and 48% (63) were level C (Figure 1). A comparison of the difficulty of conducting an RCT between these 69 level B and 63 level C

Figure 1 Algorithm for each of the recommendations in the 35 obstetric practice bulletins
Rec, recommendations; Ob, obstetrics; PB, practice bulletin; RCT, randomized clinical trial

*Considered unethical by majority of surveyors

~Considered unnecessary to do RCT by the majority of the responders

[^]Refers to practice bulletin published between June 1998 and December 2005

recommendations were notable for three factors. First, the likelihood of being categorized as least difficult (1/5) to perform was similar (12% [77/627 respondents] for level B versus 14% [76/532] for level C; OR 0.84; 95% CI 0.59, 1.18). Second, level B recommendations were significantly less likely to be the most difficult (5/5) to conduct an RCT compared with level C (19% [122/627 respondents] versus 24% [129/532]; OR 0.75; 95% CI 0.57, 0.99). Third, the majority of the respondents considered that the RCT required the MFM network similarly for levels B and C recommendations: 26% (18/69) of level B and 33% (21/63) of level C (OR 0.70; 95% CI 0.33, 1.49).

Lastly, among the 132 levels B and C recommendations that were considered ethically justifiable and in need of an RCT, seven received the highest priority to perform an RCT: two each from PB concerning premature rupture of membranes and management of preterm labour as well as one each from PB regarding shoulder dystocia, antiphospholipid antibody and perinatal viral/parasitic infections (Table 2). Most of these high priority topics (71%; 5/7) were level C recommendations. Only with two of these high priority topics did the majority of surveyors indicate that involvement of the MFM network was required for the RCT.

COMMENTS

Biased results from an inadequately designed study can mislead clinicians into making poor decisions for an individual patient and national organization to formulate erroneous guidelines.³ Since the majority of ACOG recommendations are either levels B or C,² it would be beneficial to know if these recommendations are amenable to an RCT. There are four findings of this survey.

First, 24% of levels B and C recommendations concern topics that are not amenable to an RCT because of ethical concerns by the majority of the MFM surveyed (Figure 1). Admittedly, we did not require a system of ethical criteria to judge these recommendations.⁴ We desired the clinical opinions of the

Table 1 The 16 recommendations already accepted in clinical practice as standard care

PB no.	Level B/C	ACOG PB level B or C recommendations
49	B	X-ray pelvimetry alone is as a predictor of dystocia has not been shown to have benefit and, therefore, is not recommended
36	C	Sodium citrate should be administered promptly to neutralize gastric contents following the decision to perform caesarean delivery
30	B	The screening test generally should be performed on venous plasma or serum samples using well-calibrated and well-maintained laboratory instruments
10	C	For women with third trimester fetal demise, intravaginal misoprostol may be in use
56	C	Before beginning ART, counsel women about the risk of high-order multiple gestation
19	C	Epidural anaesthesia appears to be safe in women taking unfractionated low-dose heparin if the APTT is normal
9	B	Recent normal antepartum test result should not preclude the use of intrapartum fetal heart monitoring
49	C	Women with twin gestation may undergo augmentation of labour
56	B	Higher risk of gestational diabetes and hypertension warrants screening and monitoring for these complications
24	C	Couples with otherwise unexplained RPL should be informed about the potential for successful pregnancy without treatment
29	C	Women with chronic hypertension should be evaluated for potentially reversible aetiologies, preferably prior to pregnancy
6	B	Epidural anaesthesia is safe in patients with platelet counts greater than 100,000/ μ L
10	B	The use of higher doses of misoprostol (50 μ g q 6 hours) to induce labour may be appropriate, although there are reports of increased complications like hyperstimulation
40	C	There is no evidence that any one manoeuvre is superior to another in releasing an impacted shoulder dystocia or reducing the chance of injury. However, performance of the McRoberts' manoeuvre is a reasonable initial approach
49	B	An elective cerclage can be considered in a patient with at least three unexplained mid-trimester pregnancy losses or preterm deliveries
30	C	Although universal glucose challenge screening for GDM is the most sensitive approach, there may be pregnant women at low risk* who are less likely to benefit from testing

ARI, assisted reproductive technology; GDM, gestational diabetes mellitus; PB, practice bulletin; RPL, recurrent pregnancy loss

*Low risk women should have all of the following characteristics: younger than 25 years; not a member of ethnic group with a high prevalence of diabetes (Hispanic, African, Native American, South or East Asian or Pacific Islander); body mass index of 25 or less; no history of abnormal glucose tolerance; no previous history of adverse outcomes associated with GDM; no known diabetes in first degree relative

Table 2 The top seven recommendations in need for randomized clinical trial

PB no.	Level	ACOG PB level B or C recommendations
1	C	Antenatal corticosteroids may be administered to gravidas with PROM up to 34 weeks of gestation
40	C	Planned caesarean delivery to prevent shoulder dystocia may be considered for suspected fetal macrosomia with estimated fetal weights exceeding 5000 g in women without diabetes and 4500 g in women with diabetes
1	C	Tocolytics may be utilized in patients with preterm PROM to permit administration of antenatal corticosteroids and antibiotics
20	C	Routine serological screening of all pregnant women for cytomegalovirus and toxoplasmosis is not recommended
43	B	Bed rest, hydration and pelvic rest do not appear to improve the rate of preterm birth and should not be routinely recommended
43	B	Cervical ultrasound and fetal fibronectin have good negative predictive values, thus either approach or both combined may be helpful in determining which patients do not need tocolytics
68	C	Women with antiphospholipid syndrome and no thrombotic history should receive prophylactic doses of heparin and low-dose aspirin during pregnancy and the postpartum period (6–8 weeks)

PB, practice bulletin; PROM, premature rupture of membranes

unlikely that the protocol would be approved at any institution. The second finding of the survey is that approximately one tenth (16/148) of levels B and C recommendations that are amenable to an RCT were not needed (Table 1), as they were felt to be already well entrenched in clinical practice by the vast majority of obstetricians in the USA. An examination of these 16 topics indicates that their focus was on diagnostic tests (glucose tolerance test or X-ray pelvimetry) or appropriateness of treatment (cerclage, epidural and misoprostol). This assessment may not be a universal finding; therefore, a comparison to other country's obstetrical guidelines is planned.

The third result of the survey is that of all the B- and C-level obstetric recommendations, 67% need an RCT (132/195). These 132 suggestions were equally distributed between levels B (52%) and C (48%; Figure 1). The proposed RCTs for 69 level B recommendations were significantly less likely to be rated as the most difficult to perform (5/5) compared with the 63 level C (OR 0.75; 95% CI 0.57, 0.99) and this suggests that it would be more difficult to perform an RCT on level C than that of B recommendation. However, the need for involvement of the MFM network, a surrogate for difficulty in conducting a study, was similar for the 132 levels B and C suggestions. Additionally, the lowest degree of difficulty (1/5) was similar for the levels B and C recommendations that were ethically justified and clinically needed. If the majority of RCTs are needed and can be performed, why are they not completed so that the majority of our evidence could become level A? While there are certainly more reasons than there are positions, we feel the most important involve money and time. The first issue as always, is who will pay for the RCT. Unfortunately, there is less funding for academic projects than in the past. Second, the majority of the 132 subjects that needed an RCT involve areas that are less common, thus requiring many centres or decades of study in a few locals. Alternatively, many of these subjects are very difficult to study or require long follow-ups or may, by their nature, be risky, for participants, thus making patients

respondent so that the response was based on the individuals' own ethical framework. Even if the criteria had been provided to determine if RCT is justifiable, it must be acknowledged that there is great debate about the ethics of studies involving pregnant patients and studies compliant with CONSORT guidelines.^{5–11} It can be argued whether or not a Institutional Review Board would approve of a protocol if an MFM considers the RCT to be unethical. But if the majority of respondents consider the topic not amenable to an RCT, it is

less likely to enrol. We feel the reason RCTs are not performed is unlikely to involve a lack of interest on behalf of physicians who would like to have definitive evidence on which to base their clinical decision.

The fourth focus of the survey was to identify the levels B and C recommendations that most need an RCT. From the 132 suggestions requiring an RCT, the respondents identified seven recommendations that most needed an RCT (Table 2). With two level C recommendations, 45% (5/11) of the MFMs agreed that a trial should be completed expeditiously. For two of these top seven recommendations, most of the surveyors thought the MFM network needed to undertake the study (Table 1).

There are limitations to this study and they should be acknowledged. Even though the eleven MFM were geographically dispersed throughout the country and worked in different environments, the results could have been different if more respondents were involved. However, considering we wanted to evaluate all B- and C-level recommendations involving over 900 questions, it is unlikely we could have gotten an acceptable response rate from a large number of participants. Indeed, even short questionnaires may only receive a 15–20% response rate. Another potential shortcoming of the survey was that *a priori*, we did not determine what would constitute an unethical RCT, one that was not needed, how to assess the difficulty of performing an RCT or how to identify the recommendations that urgently needed such a study. Instead, we chose to enlist reliable practitioners who had extensive experience in research, clinical care and performing RCTs. We did analyse the recommendations according to the opinion of the majority of those surveyed, as a way to triage the obstetric knowledge, and to possibly understand why the majority of recommendations are not Level A.

In summary, this data suggest that two-thirds of levels B and C recommendations are amenable and in need of an RCT.

While we used obstetric data, the principal should apply to evidence in any field of medicine.

REFERENCES

- 1 American College of Obstetricians and Gynecologists. *Reading the Medical Literature: Applying Evidence to Practice*. Washington, DC: ACOG, 1998
- 2 Chauhan SP, Berghella V, Sanderson M, Magann EF, Morrison JC. American College of Obstetricians and Gynecologists practice bulletins: An Overview. *Am J Obstet Gynecol* 2006;**194**:1564–72
- 3 Altman DG, Schulz KF, Moher D, *et al.*; CONSORT GROUP (Consolidated Standards of Reporting Trials). The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med* 2001;**134**: 663–94
- 4 McCullough LB, Coverdale JH, Chervenak FA. A comprehensive ethical framework for responsibly designing and conducting pharmacologic research that involves pregnant women. *Am J Obstet Gynecol* 2005;**193**:901–7
- 5 Stang A, Hense HW, Jockel KH, Turner EH, Tramer MR. Is it always unethical to use a placebo in a clinical trial? *PLoS Med* 2005;**2**:e72
- 6 Olson LG. Patient-centred equipoise and the ethics of randomised controlled trials. *Monash Bioeth Rev* 2002;**21**:S55–67
- 7 Kim SYH, Frank S, Holloway R, Zimmerman C, Wilson R, Kiebertz K. Science and ethics of sham surgery: a survey of Parkinson disease clinical researchers. *Arch Neurol* 2005;**62**:1357–60
- 8 Allmark P, Mason S. Should desperate volunteers be included in randomised controlled trials? *J Med Ethics* 2006;**32**:548–53
- 9 Faunce TA, Buckley NA. Of consents and CONSORTs: reporting ethics, law, and human rights in RCTs involving monitored overdose of healthy volunteers pre and post the 'CONSORT' guidelines. *J Toxicol Clin Toxicol* 2003;**41**:93–9
- 10 Djulbegovic B, Cantor A, Clarke M. The importance of preservation of the ethical principle of equipoise in the design of clinical trials: relative impact of the methodological quality domains on the treatment effect in randomized controlled trials. *Account Res* 2003;**10**:301–15
- 11 de Zulueta PE. Randomised placebo-controlled trials and HIV-infected pregnant women in developing countries. Ethical imperialism or unethical exploitation? *Bioethics* 2001;**15**:289–311

(Accepted 28 January 2009)