

Home

Contents

jameslindlibrary.org

Cautionary Tales

Grimes DA. Discovering the need for randomized controlled trials in obstetrics: a personal odyssey

Cite as: Grimes DA (2007). Discovering the need for randomized controlled trials in obstetrics: a personal odyssey. James Lind Library (www.jameslindlibrary.org).

Accessed Thursday 31 May 2007.

© David Grimes, 2007.

Author contact details: David Grimes, Family Health International, PO Box 13950, Research Triangle Park, North Carolina 27709, USA. E-mail: dgrimes@fhi.org

A naïve and impressionable medical student, I found the world of obstetrics in the early 1970's to be strange and wonderful. Like others (King 2005), only after completing my clinical training and later being introduced to critical appraisal was I to learn just how strange it really was. The grand mystery of bringing forth new life was shrouded in arcane medical and nursing-midwifery rituals (Klein et al. 2006). Some of these venerable traditions were not only inappropriate but also harmful. These customs stemmed from a number of sources, including seduction by authority, the false idol of technology, the tendency to let sleeping dogmas lie, and the pursuit of pedantry in medicine and nursing-midwifery (Grimes 1986). I will provide examples of each from my obstetrical odyssey.

Seduction by authority

The epidemic of diethylstilbestrol (DES) harms unfolded during my early years in medicine. As a trainee, I began to see patients with a bizarre "cockscorn" configuration (Figure 1) of the visible cervix....a common stigma among "DES daughters" (women exposed in utero to DES.)

Based largely on the advocacy of Drs George and Olive Smith of Boston (Smith et al. 1946), an estimated 5 to 10 million US women were prescribed DES during pregnancy. The Smiths argued that DES administration promoted production of progesterone by the placenta, which allegedly had a salutary effect on pregnancy. Among the putative benefits were reductions in the risk of threatened miscarriage, recurrent miscarriage, and premature labor. Because the drug did not have patent protection, several hundred companies in the U.S. manufactured or distributed DES, extolling the health benefits of this "wonder drug" (Figure 2).

The eminence of the Boston physicians trumped the evidence (Isaacs and Fitzgerald 1999). A placebo-controlled randomized controlled trial published by Dieckmann and others (1953) failed to find any benefit. Despite the lack of solid evidence that DES prevented adverse outcomes, wide use continued. Decades later, a landmark case-control study highlighted the dangers of fetal exposure to DES, especially clear cell carcinoma of the vagina in the daughters of women who had received the drug during pregnancy (Herbst et al. 1971). Not until 1971 did the U. S. Food and Drug Administration first warn against use of DES in pregnancy. By the mid-1970's, more than 200 cancers of the vagina, most associated with DES exposure, had been reported to a DES registry (Herbst et al. 1977). Because of the prominence of the Boston authorities, credible data were ignored because the problem was serious and DES "might work."

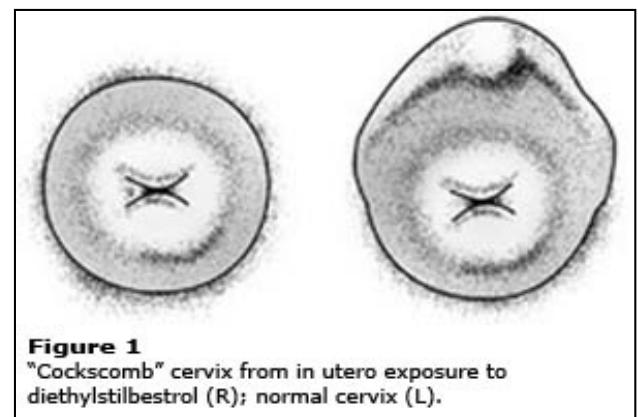
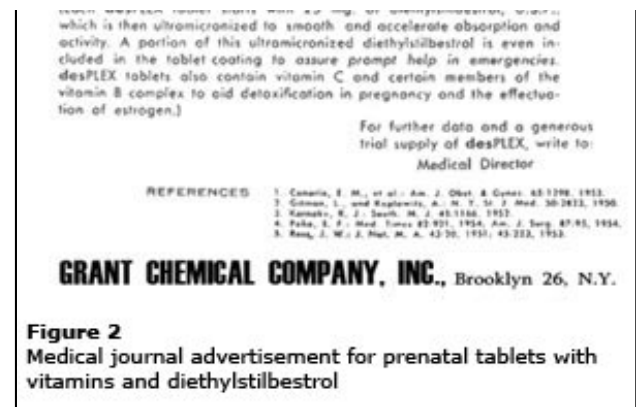


Figure 1
"Cockscorn" cervix from in utero exposure to diethylstilbestrol (R); normal cervix (L).



The false idol of technology

I remember clearly the morning when the first electronic fetal monitor was wheeled onto our labor and delivery suite by the smiling salesman. Our hopes were as high as his, since our professors had assured us this would be our "window into the womb." By beaming ultrasound continuously into the uterus (or by screwing an electrode into the fetal scalp), we could track the heart rate of the fetus and relate it to uterine contractions. Replete with blinking lights, colorful paper tracings, abdominal belts, and tangles of wires, electronic fetal monitoring took the country by storm ([Figure 3](#)).

The rise of fetal monitoring in the 1970's paralleled the uncritical use of DES three decades earlier, based largely on the advocacy of a few prominent obstetricians at well-known medical institutions. The scientific evidence supporting its use was meager: before-after studies showing lower perinatal mortality rates in the era of electronic fetal monitoring than in earlier times. Regrettably, the advocates of this technology were unaware of screening test principles: even a good test (which electronic fetal monitoring is not) used in a setting with a low prevalence of disease (for example, fetal death during labor) will have a dismal positive predictive value. Positive tests are almost always wrong.

The response to this new technology was predictable: it quickly became a multibillion-dollar industry, and most births today in the US involve electronic fetal monitoring. Equally predictable was the resultant increase in operative deliveries, especially cesarean birth. When we clinicians are unsure of what the tracing means, the "safest" course is to intervene. The U.S. epidemic of cesarean deliveries has been fueled in part by over-reaction to false-positive traces from electronic fetal monitoring.



By the time randomized controlled trials were finally done, the practice had become entrenched. Electronic fetal monitoring had become the de facto standard of practice....not because it was better than a nurse-midwife listening to the fetus with a stethoscope but because it was cheaper. The monitor was a one-time purchase and required few supplies, while nurse-midwife salaries are the largest component of most hospital budgets. Given the choice between a gadget or a nurse-midwife at the bedside, technology prevailed.

Electronic fetal monitoring failed to fulfill its bright promise. As revealed in a recent Cochrane review (Alfirevic et al. 2006), the use of operative delivery, especially cesarean birth, increased by two-thirds (RR 1.66; 95% CI 1.30-2.13), but no lasting benefits were evident for the fetus or baby. Although the risk of neonatal seizures was reduced by half (RR 0.50; 95% CI 0.31-0.80), the risk of cerebral palsy was paradoxically increased, although the difference was not statistically significant (RR 1.74; 95% CI 0.97-3.11). In brief, electronic fetal monitoring increased risk to the mothers with no evidence of lasting benefit to their babies.

Nonetheless, electronic fetal monitoring remains entrenched. Major academic centers sponsor courses on interpretation of heart-rate patterns, despite the lack of uniform definitions after three decades of use (Parer and Kind 2000). In the most charitable assessment to date, the positive predictive value of electronic fetal monitoring for acidemia at birth was 0.37 (Vintzileos et al. 1995). Stated alternatively, when the heart-rate tracing looks worrisome, the fetus has a normal blood pH two times out of three. This positive predictive value is lower than that of flipping a coin (0.50). The coin is smaller, simpler, and cheaper than an electronic fetal monitor.

We have learned from our monitoring mistakes, however. In the 1990's, home uterine activity monitoring threatened to become another juggernaut. Premature birth remains the greatest challenge in obstetrics, and the incidence has been increasing in recent years. With home uterine activity monitoring, a woman deemed to be at increased risk of preterm labor applies an external contraction monitor to her abdomen twice daily at home. The results are then transmitted by telemetry to a clinician, who reviews the tracings to look for incipient labor. If early labor is identified, the woman seeks obstetrical care -- presumably earlier than would have occurred in the absence of the home monitoring.

An early industry-sponsored randomized controlled trial, improperly analyzed, reported a benefit (Hill et al. 1990). Another trial (having failed to find any overall benefit) touted improvements in the subgroup of women with twins (Dyson et al. 1991). Concerned about this misinterpretation, we did a systematic review of all the evidence and were unable to detect any benefit (Grimes and Schulz 1992). Soon thereafter, ingenious randomized controlled trials with sham monitoring arms confirmed that home uterine activity monitoring does not improve outcomes, despite its large costs (Dyson et al. 1998; Collaborative Home Uterine Monitoring Study Group 1995). In contrast to electronic fetal heart rate monitoring during labour, home uterine monitoring during pregnancy is now disappearing from practice.

Letting sleeping dogmas lie

Inertia is the great flywheel of society – and of medical and nursing-midwifery practice as well. As a budding clinician, I was taught to perform routine elective episiotomy for women giving birth the first time; this would prevent a nasty spontaneous tear, which might extend into the rectum. This sage advice came from revered figures in American obstetrics: Drs Pomeroy and DeLee. The latter argued for episiotomy because he believed that vaginal birth was “decidedly pathological”tantamount to impalement on a pitchfork (DeLee 1920). Clinicians were duty-bound to intervene. By the 1980's, more than 60% of all US births (and the large majority of first births) had this operation, making it one of the most common surgical procedures performed in the twentieth century. After decades of unquestioned practice, more recent evidence from systematic reviews of controlled trials has shown that episiotomy should be restricted (Klein et al. 2006; Rooks 1999). Indeed, midline episiotomy paradoxically increases the risk of an extension of the defect into the rectum. In response to this evidence, episiotomy rates in the US have continued to drop (Rooks et al. 1999).

During my training, 24-hour monitoring of urine estriol excretion was de rigueur for pregnancies thought to be in jeopardy. The discovery that this estrogen is a unique product of the fetus and placenta led to widespread estriol screening to monitor fetal well-being. Women collected all their urine over a day and dutifully toted plastic jugs of urine to the clinic. Some came to clinic with their bottles concealed in paper bags for discretion; others described their embarrassment in keeping urine collections in the family refrigerator. When the estriol content was found to be abnormally low (implying a fetus in peril), we usually ignored it and attributed the decline to an incomplete urine collection. Concerned about the expense, inconvenience, and indignity for our patients, disgruntled residents at my hospital reviewed the literature and suggested that we abandon the test. Our professor replied, “We’re a university hospital; we have to provide this test.” (We had thought a university hospital just might be the first to give it up.) Ultimately, a randomized controlled trial was done, which failed to demonstrate any significant benefit (Neilson and Cloherty 2000). Twenty-four-hour urine collections disappeared from obstetrical practice...and from family refrigerators.

Uncritical acceptance of new technologies has been a stubborn problem in obstetrics. After estriol collection was abandoned, its successor was arguably worse: antepartum assessments of fetal heart-rate patterns. With this test, women thought to be at increased risk of poor outcomes come to the clinic frequently for recording of the fetal heart; these patterns are then correlated with spontaneous or induced uterine contractions. Paradoxically, the six randomized controlled trials of this technology, though now old, offer no evidence of benefit (Pattison and Cowan 2000). Indeed, they suggest an increased risk of perinatal death when antepartum cardiotocography is used. Predictably, this has done little to discourage its use.

Intravenous magnesium sulfate (Epsom salt) to stop premature labor is a North American anomaly (Figure 4). Enthusiasm for the practice stems from one poorly reported trial with an improper outcome and a couple of reports of uncontrolled case-series(Grimes and Nanda 2006). There is no evidence from randomized controlled trials that magnesium sulfate is effective as a tocolytic agent (Crowther et al. 2002), but calcium channel blockers, such as nifedipine, seem promising (King et al. 2003).



When we encouraged our colleagues to abandon magnesium sulfate in favor of calcium channel blockers for tocolysis, however, our commentary (Grimes and Nanda 2006) caused a stir. One perinatologist urged continued use of magnesium sulfate until something better came along (Perkins 2007), despite evidence that something better had indeed come along.



Figure 4
Epsom salt (magnesium sulfate), an agent with many putative uses

The pursuit of pedantry

Early in my training, X-ray pelvimetry (calculation of several diameters of the maternal bony pelvis) was an arcane quantitative science. In that era, performance of a cesarean delivery was an admission of obstetrical defeat. X-ray pelvimetry was required before a cesarean delivery could be entertained or performed. Women - often in booming labor - had to be transported by stretcher over long distances in hospitals to the radiology department. The most junior member of the obstetrical team (often me) accompanied the woman, along with an emergency delivery kit (just in case). Radiology staff dreaded X-ray pelvimetry; fear of a delivery occurring in the radiology department was palpable. Apocryphal tales of deliveries on the X-ray table, in hospital corridors, or in elevators were part of our hospital lore, and the oxytocic effects of horizontal travel by stretcher (or perhaps just the change of scenery) were legendary.

Except for rachitic dwarfs (rare then and now), the X-ray pelvimetry report was usually inconclusive. We seldom acted on the measurements; rather, we allowed the labor to continue independent of the pelvimetry results. The adequacy of the pelvis was ultimately decided by the progress of labor (or lack thereof). Measuring the various pelvic diameters struck me as a unrewarding exercise ... a static, two-dimensional assessments of the bony pelvis. These provided no insight into molding of the fetal skull, separation of the symphysis, or contractile forces involved.

My dim view of pelvimetry proved correct. The available randomized controlled trials of X-ray pelvimetry, though of limited quality, found that it increased the frequency of cesarean delivery without any detectable benefit to the baby (a familiar theme) (Pattinson 2000). The predictive values of positive and negative tests were similar to flipping a coin (a familiar theme). In addition to the inconvenience and expense of these radiographs, evidence now suggests that this in utero exposure of fetuses to radiation increases the risk of cancer in later life (a familiar theme) (Schussman and Lutz 1982).

Pedantry extended to the nursing-midwifery staff, who ran the labor and delivery suite. An immutable rule at my hospital was that no woman was to deliver without a perineal shave and a one-liter enema. The shaving seemed barbaric (or barber-ic?), especially for women in active labor. Plus, the putative link between pubic hair and obstetrical morbidity eluded me then, and today (Hofmeyr 2005). Indeed, as long ago as 1922, Johnson and Sidall challenged this notion in one of the earliest reports of a controlled trial, and were unable to find any support for the practice ([Johnson and Sidall 1922](#)).

The routine enema policy was even more bizarre. Some feces might pass in the second stage of labor without an enema; however, forcible ejection of residual enema fluid could also occur, as my gowns occasionally testified. For women in late labor, it was sometimes a race to see which exited first: fetus or enema. Occasionally, it was a tie, to everyone's dismay.

These archaic, embarrassing practices have now disappeared, since evidence of benefit remains lacking (Cuervo et al. 2000; Basevi and Lavender 2001).

From wooden to silver spoon

A sea change occurred in obstetrical practice during my career. Having been awarded a "wooden spoon" ([Figure 5](#)) by Archie Cochrane in 1979 for the worst use of randomized controlled trials, obstetrics promptly transformed itself (King 2005).

Obstetricians in the UK, The Netherlands, and Canada led an international effort to find the best available evidence to guide clinical practice and identify, catalogue, and disseminate the



Figure 5
Last wooden spoon awarded at St. Johns College, Cambridge University, 1909. A British booby prize,

results of the world's randomized controlled trials in obstetrics and perinatology. This culminated in the 1989 publication of *Effective Care in Pregnancy and Childbirth*, the first evidence-based obstetric text, and *The Oxford Database of Perinatal Trials*, an electronic publication that allowed the analyses published in the book to be updated as new evidence became available or mistakes identified. This international collaboration, which was coordinated from the National Perinatal Epidemiology Unit in Oxford, provided the model for the Cochrane Collaboration (www.cochrane.org), which has extended this mission to other areas of health care. Within the span of a decade, obstetrics moved from the scientific backwaters to scientific prominence. Indeed, in his foreword to *Effective Care in Pregnancy and Childbirth*, Archie Cochrane happily rescinded his "wooden spoon" award, noting that "I now have no hesitation whatsoever in withdrawing the slur of the wooden spoon from obstetrics, and I feel honoured by being associated, even in an indirect way, with such an important publication." (Cochrane 1989).

Last wooden spoon awarded at St. Johns College, Cambridge University, 1909. A British booby prize, it was given annually to the student with the lowest score on the comprehensive mathematics exam at Cambridge University -- implying that the student was better equipped to be a cook than a scholar.

Obstetrical practice today is increasingly guided by high-quality evidence from randomized controlled trials. Examples include administration of corticosteroids to women at risk of preterm delivery (Roberts and Dalziel 2006), provision of high-dose folate to women who have had a fetus with neural tube defect (Lumley et al. 2001), and active management of the third stage of labor (Prendiville et al. 2000). Evidence-based medicine has gained legitimacy in obstetrics, and practice guidelines of professional organizations such as the Royal College of Obstetricians and Gynaecologists and American College of Obstetricians and Gynecologists increasingly reflect the best available evidence.

Uptake of evidence from randomized controlled trials has not been uniform, however. Audits in different parts of the world reveal dramatically different penetration of evidence-based medicine into hospital practice (Grimes 1995). In a tertiary-care hospital in Birmingham, UK, in 1998-1999, 325 consecutive obstetrical and gynecological admissions were scrutinized for the evidence supporting the interventions provided. A large majority (90%) had "substantial research evidence" for the interventions (Khan et al. 2006).

In contrast, an audit of obstetrical practices at a large Egyptian maternity hospital found many inconsistencies with evidence-based practice recommendations. When observed practices were compared with the 1999 World Health Organization classification of practices for normal birth, some beneficial practices were infrequent, while other harmful practices were common. For example, oxytocin was given prophylactically after delivery to only 15% of women. In contrast, intravenous oxytocin infusions during labor were administered inappropriately to 93% of women, and these infusions were commonly unlabeled and not monitored (Khalil et al. 2005). To help remedy these disparities, the World Health Organization's Reproductive Health Library disseminates relevant Cochrane reviews to clinicians worldwide, although evidence of impact on practice is limited to date (Gülmezoğlu et al. 2007).

Archie Cochrane would be proud of obstetrics today. The influence of expert opinion, new technologies, tradition, and pedantry (Grimes 1986) is being supplanted by systematic reviews of the best available evidence (Grimshaw 2004). Despite substantial progress, large disparities persist between what we know and what we do. To paraphrase Bertrand Russell, excellent care should be inspired by compassion and guided by science. Randomized controlled trials - and especially systematic reviews of them (Cook et al. 1997) - help ensure that our obstetrical care is as well-guided as it is well-intended.

References

- Alfirevic Z, Devane D, Gyte GM (2006). Continuous cardiotocography (CTG) as a form of electronic fetal monitoring (EFM) for fetal assessment during labour. *Cochrane Database Syst Rev* 3:CD006066.
- Basevi V, Lavender T (2001). Routine perineal shaving on admission in labour. *Cochrane Database Syst Rev* CD001236.
- Cochrane AL (1989). Foreword. In: Chalmers I, Enkin M, Keirse MJNC, eds. *Effective care in pregnancy and childbirth*. Oxford: Oxford University Press.
- Collaborative Home Uterine Monitoring Study (CHUMS) Group (1995). A multicenter randomized controlled trial of home uterine monitoring: active versus sham device. *Am J Obstet Gynecol* 173:1120-7.
- Cook DJ, Mulrow CD, Haynes RB (1997). Systematic reviews: synthesis of best evidence for clinical decisions. *Ann Intern Med* 126:376-80.
- Crowther CA, Hiller JE, Doyle LW (2002). Magnesium sulphate for preventing preterm birth in threatened preterm labour. *Cochrane Database Syst Rev* CD001060.

- Cuervo LG, Rodriguez MN, Delgado MB (2000). Enemas during labor. *Cochrane Database Syst Rev* CD000330.
- DeLee JB (1920). The prophylactic forceps operation. *Am J Obstet Gynecol* 1:34-44.
- Dieckmann WJ, Davis ME, Rynkiewicz LM, Pottinger RE (1953). Does the administration of diethylstilbestrol during pregnancy have therapeutic value? *Am J Obstet Gynecol* 66:1062-81.
- Dyson DC, Crites YM, Ray DA, Armstrong MA (1991). Prevention of preterm birth in high-risk patients: the role of education and provider contact versus home uterine monitoring. *Am J Obstet Gynecol* 164:756-62.
- Dyson DC, Danbe KH, Bamber JA, Crites YM, Field DR, Maier JA, et al (1998). Monitoring women at risk for preterm labor. *N Engl J Med* 338:15-9.
- Grimes DA (1986). How can we translate good science into good perinatal care? *Birth* 13:83-90.
- Grimes DA (1995). Introducing evidence-based medicine into a department of obstetrics and gynecology. *Obstet Gynecol* 86:451-7.
- Grimes DA, Nanda K (2006). Magnesium sulfate tocolysis: time to quit. *Obstet Gynecol* 108:986-9.
- Grimes DA, Schulz KF (1992). Randomized controlled trials of home uterine activity monitoring: a review and critique. *Obstet Gynecol* 79:137-42.
- Grimshaw J (2004). So what has the Cochrane Collaboration ever done for us? A report card on the first 10 years. *CMAJ* 171:747-9.
- Gülmezoğlu AM, Langer A, Piaggio G, Lumbiganon P, Villar J, Grimshaw J (2007). Cluster randomised trial of an active, multifaceted educational intervention based on the WHO Reproductive Health Library to improve obstetric practices. *BJOG* 114:16-23.
- Herbst AL, Ulfelder H, Poskanzer DC (1971). Adenocarcinoma of the vagina. Association of maternal stilbestrol therapy with tumor appearance in young women. *N Engl J Med* 284:878-81.
- Herbst AL, Cole P, Colton T, Robboy SJ, Scully RE (1977). Age-incidence and risk of diethylstilbestrol-related clear cell adenocarcinoma of the vagina and cervix. *Am J Obstet Gynecol* 128:43-50.
- Hill WC, Fleming AD, Martin RW, Hamer C, Knuppel RA, Lake MF, et al (1990). Home uterine activity monitoring is associated with a reduction in preterm birth. *Obstet Gynecol* 76:13S-8S.
- Hofmeyr GJ (2005). Evidence-based intrapartum care. *Best Pract Res Clin Obstet Gynaecol* 19:103-15.
- Isaacs D, Fitzgerald D (1999). Seven alternatives to evidence based medicine. *BMJ* 319:1618.
- Johnston RA, Sidall RS (1922).** Is the usual method of preparing patients for delivery beneficial or necessary? *Am J Obstet Gynecol* 4:645-650.
- King JF (2005). A short history of evidence-based obstetric care. *Best Pract Res Clin Obstet Gynaecol* 19:3-14.
- Khalil K, Elnourey A, Cherine M, Sholkamy H, Hassanein N, Mohsen L, et al. (2005) Hospital practice versus evidence-based obstetrics: categorizing practices for normal birth in an Egyptian teaching hospital. *Birth* 32:283-90.
- Khan AT, Mehr MN, Gaynor AM, Bowcock M, Khan KS (2006). Is general inpatient obstetrics and gynaecology evidence-based? A survey of practice with critical review of methodological issues. *BMC Womens Health* 6:5.
- King JF, Flenady VJ, Papatsonis DN, Dekker GA, Carbone B (2003). Calcium channel blockers for inhibiting preterm labour. *Cochrane Database Syst Rev* CD002255.
- Klein MC, Sakala C, Simkin P, Davis-Floyd R, Rooks JP, Pincus J (2006). Why do women go along with this stuff? *Birth* 33:245-50.

Lumley J, Watson L, Watson M, Bower C (2001). Periconceptional supplementation with folate and/or multivitamins for preventing neural tube defects. Cochrane Database Syst Rev CD001056.

Neilson JP, Cloherty LJ (2000). Hormonal placental function tests for fetal assessment in high risk pregnancies. Cochrane Database Syst Rev CD000108.

Parer JT, King T (2000). Fetal heart rate monitoring: is it salvageable? Am J Obstet Gynecol 182:982-7.

Pattinson RC (2000). Pelvimetry for fetal cephalic presentations at term. Cochrane Database Syst Rev CD000161.

Pattison N, McCowan L (2000). Cardiotocography for antepartum fetal assessment. Cochrane Database Syst Rev CD001068.

Perkins RP (2007). Magnesium sulfate tocolysis: time to quit. Obstet Gynecol 109:778-9; author reply 779.

Prendiville WJ, Elbourne D, McDonald S (2000). Active versus expectant management in the third stage of labour. Cochrane Database Syst Rev CD000007.

Roberts D, Dalziel S (2006). Antenatal corticosteroids for accelerating fetal lung maturation for women at risk of preterm birth. Cochrane Database Syst Rev 3:CD004454.

Rooks JP (1999). Evidence-based practice and its application to childbirth care for low-risk women. J Nurse Midwifery 44:355-69.

Schussman LC, Lutz LJ (1982). Hazards and uses of prenatal diagnostic X-radiation. J Fam Pract 14:473-80.

Smith OW, Smith GVS, Hurwitz D (1946). Increased excretion of pregnanediol in pregnancy from diethylstilbestrol with special reference to the prevention of late pregnancy accidents. Am J Obstet Gynecol 51:411-415.

Vintzileos AM, Nochimson DJ, Antsaklis A, Varvarigos I, Guzman ER, Knuppel RA (1995). Comparison of intrapartum electronic fetal heart rate monitoring versus intermittent auscultation in detecting fetal acidemia at birth. Am J Obstet Gynecol 173:1021-4.

[Home](#)[Contents](#)

[Comments welcome](#)