Agency for Healthcare Research and Quality

Evidence Report/Technology Assessment

Number 71

Vaginal Birth After Cesarean (VBAC)

Summary

Purpose of Report and Target Audience

This report provides a framework for comparing the harms and benefits of delivery options for women with prior cesarean delivery (CD). The information is designed to help consumers, providers, payers, and policymakers in decisionmaking about repeat cesarean or trial of labor (TOL).

Overview

In 2000, 22.9 percent of all births in the United States occurred by CD. This rate is the highest total CD rate reported since data collection began in 1989. The vaginal birth after cesarean (VBAC) rate, defined as the proportion of women with a prior CD who delivered vaginally, steadily increased from 1989 to 1996. As allowing TOL became more common, practice variation became a larger concern, e.g., expanding criteria for eligibility and medical induction, and for augmentation of labor. In parallel with this liberalization of criteria and management, highly publicized articles suggested that maternal and fetal risks were perceived to be increasing. Subsequently, the VBAC rate has decreased 27 percent from 1996 to 2000. Currently, a crisis in malpractice rates is decreasing the availability of maternity care providers and raising concerns that patients may have limited options, less access to care, and perhaps be at increased risk for complications.

Reporting the Evidence

The strength and suitability of the evidence regarding the risks of major maternal and infant morbidity and mortality associated with TOL or elective repeat cesarean delivery (ERCD) in women with prior low transverse of unknown scar. The scope of the review was to examine events that were specifically related to having had a prior CD. Comparisons purely about vaginal versus cesarean delivery such as incontinence, pelvic support disorders, and respiratory consequences but not specifically about VBAC or repeat cesarean, were not considered, though these topics are important to consider when deciding upon route of delivery. In judging the suitability of evidence, we took the perspective that the first thing a decision-maker would want to know is whether the risk of these complications is higher for a trial of labor, versus an elective cesarean delivery, under optimal conditions of care. That is, the most relevant evidence would compare the outcomes and risks of a properly managed trial of labor to that of a properly conducted elective cesarean delivery. Some components of obstetric care, as well as some aspects of the setting of this care, might increase the risks of TOL or ERCD. For example, it has been hypothesized that the use (or misuse) of drugs for induction and augmentation might increase the risk of uterine rupture in patients who have had a prior cesarean delivery. We examined the strength of evidence that these factors influence these outcomes and adverse effects and to what extent these factors can explain the results of observational studies of VBAC complications.

Methodology

Key Questions

Two types of key questions were addressed. The first group (Questions 1- 7) compares the outcomes of a TOL and an ERCD:

1. What is the frequency of vaginal delivery in women who undergo a TOL (spontaneous onset, induced, and augmented) after prior low transverse cesarean or unknown scar?
2. How accurate are risk assessment tools for identifying patients who will have a vaginal delivery after a TOL?
3. What are the relative harms associated with a TOL (spontaneous onset, induced, and augmented) and repeat cesarean?
4. What is the incidence of uterine rupture, and are there methods for preventing major morbidity and mortality due to uterine rupture?
5. What are the health status and health-related quality of life for VBAC and repeat cesarean patients?
6. Regarding VBAC and repeat cesarean, what factors influence patient satisfaction/dissatisfaction with their childbirth experience?
7. How are economic outcomes related to VBAC, repeat CD, and their respective complications?

The second group (Questions 8-10) address factors influencing the decision to have a TOL:

8. What individual factors influence route of delivery?
9. What factors influence a patient's decisionmaking regarding VBAC or ERCD?
10. How do legislation, policy, guidelines, provider characteristics, insurance type, and access to care affect health outcomes for VBAC candidates?

Relevant studies were identified from multiple searches of MEDLINE® (1966 to 2002) and HealthSTAR (1975 to 2002), from the reference lists of systematic reviews and from local and national experts. The online Cochrane systematic reviews and controlled trials registries, DARE, National Centre for Reviews and Dissemination, and EMBASE databases were searched for relevant literature on specific topics as well. For topics related to patient preferences and satisfaction, PsycINFO and CINAHL® databases were searched. Databases were searched twice during the course of the project, with the final search in March 2002. For all VBAC topics combined, 14,449 citations were retrieved, including 4,867 about spontaneous labor and uterine rupture, 2,528 about ERCD, 2,416 about induction of labor, 2,945 citations about predictors, 1,257 about patient satisfaction, preference and health status, and 436 about cost and access.

All searches were limited to English-language articles published since 1980 (the date of the NIH Consensus Conference on VBAC) in developed countries. The report focused on studies that identified a group of patients with prior cesarean. For patient preferences and satisfaction, studies of the general birthing population, were considered if there were no studies that identified patients with prior cesarean. Studies were excluded if they focused on patients with particular conditions such as gestational diabetes, HIV, preeclampsia, and so on. Exclusions were also made for studies that focused primarily on the following: nulliparous women, vertical, lower vertical, "classical" or "classic" cesarean, vaginal breech delivery, preterm delivery, multiple gestation, or low birth weight.

Two investigators reviewed a random set of titles and abstracts for each topic to select articles for full-text review. When an appropriate level of reliability was reached for inclusion and exclusion of studies, the primary investigator reviewed the remaining titles and abstracts on the topic. Investigators read the full-text version of the retrieved papers and reapplied the initial eligibility criteria. Data from 224 studies were abstracted and included in the evidence tables described in the results section of this report.

Data Abstraction

Included study designs were determined by topic area. Study designs of included articles consisted of randomized controlled trials, cohort studies, case-control studies, cross-sectional studies, large case series (more than 10 subjects), and economic or decision models. All data were abstracted by the lead investigator for the topic. If the lead investigator encountered difficulty in finding or interpreting information in the published report, a second investigator reviewed the article and a consensus was reached.

Assessment of Study Quality

To assess the internal validity of individual studies, we applied a set of design-specific criteria developed by the current U.S. Preventive Services Task Force and additional criteria developed by the NHS Centre for Reviews and Dissemination, based at the University of York in England. In general, studies were rated good if they met all criteria, fair if they addressed some but not all criteria, and poor if they had a "fatal flaw." Investigators were asked to use the study quality ratings as previously described to determine for their topic which quality components were most important in assessing internal validity. This process allowed for some individual topic fit for fatal flaws, etc. A second investigator independently rated all included articles, and disagreements were resolved by consensus.

Data Synthesis

Where appropriate, meta-analysis was performed using WinBugs® or StatsDirect® software. To reduce potential bias, only studies of fair or good quality were included in the analyses.

Findings

Question 1. Likelihood of Vaginal Delivery

- Rates of vaginal delivery when attempting TOL ranged from 60 to 82 percent. The largest population-based

study reported a rate of 60.4 percent. The combined vaginal delivery rate for all prospective cohort studies, largely conducted in tertiary care centers and University settings, was 75.9 percent.

- There are limited data on the effect of medical induction and augmentation of labor.
- There was a 10-percent reduction in the likelihood of vaginal delivery when oxytocin was used for ether induction or augmentation. There was a similar trend in reduced likelihood of vaginal delivery with prostaglandins.

Question 2. Predictive Tools

- Two validated scoring systems categorized women into groups with likelihoods of vaginal delivery ranging from roughly 45 to 95 percent.
- One tool was able to stratify more of the population (up to 50 percent of women choosing TOL) into high and low probability subgroups, with a relatively low false-positive rate.
- By using a prospective cohort design and the largest study population, the best scoring system created a 10-point score based on the presence or absence of five variables commonly available for most patient admissions.
- An RCT clearly demonstrated the inability of X-ray pelvimetry (XRP) to predict route of delivery reliably.
- Imaging studies that combined the measurements of the pelvis and fetus showed promising results, but were limited by their lack of control for confounding and biases.

Question 3. Maternal and Infant Outcomes

General

- In the absence of RCTs of TOL versus repeat cesarean, evidence that is most generalizable comes from large country, State, or regional population-based studies (referred to as population-based studies) followed by large multicenter cohort studies, large single-institution or single-practice cohort studies, then smaller cohort studies, respectively.
- There is no direct evidence regarding the benefits and harms of TOL relative to ECRD in women who are similar in every respect except choice of delivery route.
- Several fair and good quality studies provide indirect evidence about relative benefits and harms of each route.

Maternal

- Maternal death rates did not differ between TOL and ERCD.
- The best evidence suggests that hysterectomy rates do not differ between TOL and ERCD.
- No studies examined specifically the risks of incontinence or pelvic support disorders in women with prior cesarean.
- Rates of infection were increased in ERCD versus TOL overall. Studies that performed subgroup analyses for TOL with and without vaginal delivery consistently found increased rates of infection for women who attempted TOL but ultimately had a cesarean delivery.
- There is conflicting evidence regarding whether induction of labor affects infection rates.

Infant

- There is insufficient evidence regarding the effect of selected route of delivery and Apgar score or respiratory morbidity.
- No study measured infant death directly attributable to a mother's choice of TOL or repeat CD.
- There is uncertainty about the magnitude of risk of perinatal death due to TOL. Results from two large studies differ in the magnitude of increased risk from TOL versus ERCD (90/1,000 TOL versus 50/1,000 ERCD compared with 12.9/1,000 TOL versus 1.1/1,000 ERCD). Neither study provides direct evidence of risk.

Question 4. Uterine Rupture

- The use of terms among studies is inconsistent.
- Definitions among studies for similar terms are ambiguous.
- There is no difference in asymptomatic uterine rupture rates in TOL versus ERCD.
- Symptomatic uterine rupture is significantly more common in TOL versus ERCD, with an increased risk of 2.7/1000.
- Based on the frequency and severity of symptomatic uterine rupture, the risk of perinatal death due to a rupture of a uterine scar is 1.5/10,000 and the risk of maternal hysterectomy is 4.8/10,000. These rates of serious complications such as perinatal death are probably more precise than overall risks from studies measuring death directly.
- The definition of uterine rupture as an outcome is confounded by a definition that includes the potential predictor of fetal heart rate (FHR) tracing abnormality.
- Measurement of frequency of occurrence, predictors for what population is at greatest risk, and predictors for poor outcomes are not possible, because of the lack of standard case definition.

Question 5. Health Status

- There were no studies of health status or health-related quality of life for VBAC or repeat CD patients.

Question 6. Patient Satisfaction

- Studies of patient satisfaction largely consisted of the patient's own provider obtaining information about patient satisfaction, introducing the possibility of measurement bias.
- Only two cross-sectional studies used methods other than the patient's own provider to obtain satisfaction information.
- No study measured satisfaction for the three types of delivery outcomes that could be experienced by women with prior CDs (VBAC, TOL followed by CD, or ERCD).

Question 7. Cost and Health Care Resources

- For a TOL success probability of 76 percent or greater, TOL is more cost-effective and provides higher quality of life.
- Further evaluation is needed of the sensitivity of the probability cut point of 76 percent to other potential predictor variables.

Question 8. Individual Factors

- The vast majority of studies looking at individual factors that influence the route of delivery were of poor quality due to the lack of control for confounding factors.
- The factors that were significantly associated with an increased likelihood of vaginal delivery (i.e., successful TOL) were maternal age less than 40 years, prior vaginal delivery (particularly vaginal delivery after cesarean), a nonrecurrent indication for the prior CD, and favorable cervical factors.
- The factors that were significantly associated with a decreased likelihood of vaginal delivery (i.e., failed TOL) were an increasing number of prior CD, gestational age greater than 40 weeks, birthweight greater than 4000 g, and augmentation of labor.

Question 9. Patient Preferences

- Patient preferences for birth choice are unclear because of the heterogeneity of the 11 included studies.
- Several factors appear related to choice for TOL (White race, prior vaginal delivery, lower levels of anxiety during the pregnancy).
- Lack of medical information along with cultural ideologies might account for minority women being less likely to attempt a TOL when compared with White women.
- A woman's choice for delivery was often based on social motives (e.g., easier recovery, so she can care for baby and children at home).
- Only four of 11 studies cited safety for mother or baby as important reasons for delivery choice.
- It remains unclear whether VBAC education increases the proportion of women who choose TOL.

Question 10. Legal, Provider, Hospital, Insurance Characteristics

General

- Studies of legislation, policy, guidelines, hospital characteristics, provider characteristics, insurance type, or access to care focus exclusively on VBAC rates rather than safety.

Legal

- No study provides direct evidence for the impact of malpractice issues on VBAC or ERCD.
- One study reported that VBAC rates increased when legislation was enacted that standardized VBAC guidelines had to be provided to obstetric providers.
- The best evidence suggests that use of opinion leaders provides a greater likelihood of changing practice compared with audit and feedback.

Provider

- Studies of provider characteristics failed to control for important variables such as patient selection bias.

Hospital

- VBAC rates were higher in teaching hospitals compared to private, community, regional, or non-teaching hospitals.
- Three studies conflicted over the effect of hospitals containing a neonatal intensive care unit (NICU).

Insurance

- There was conflicting evidence regarding whether insurance status predicts VBAC.

Summary of Evidence

The following summarizes the type of study design, the quality of the evidence from studies, and the suitability of the study design to answer the particular question for each key question.

Summary of Evidence of Key Questions

Key Question	Study Type*	Quality of Evidence	Suitability of Study Design†
Question 1			
What is the frequency of vaginal delivery in women who undergo a TOL (spontaneous onset, induced, and augmented) after prior low transverse cesarean or unknown scar?	II-2	Fair-Good: Several large prospective and retrospective studies; mostly consistent findings.	Greatest
Question 2			
How accurate are risk assessment tools for identifying patients who will have a vaginal delivery after a TOL?			
Predictive tools	II-2	Fair-Good: Large fair and good quality cohort studies suggest tools can provide additional information to predict likelihood of vaginal delivery.	Greatest
Imaging modalities	I	Good: Good quality RCT demonstrated that imaging was ineffective to predict vaginal birth.	Greatest
Question 3			
What are the relative harms associated with a TOL (spontaneous onset, induced and augmented) and repeat cesarean?	II-2	Fair-Poor: Several large cohort studies were inconsistent in their definitions for important health outcomes.	Moderate
Maternal Death		Fair: Studies consistently found no increased risk of maternal death from TOL versus ERCD.	Least
Hysterectomy		Fair-Poor: Many studies failed to report indication for hysterectomy.	Moderate
Transfusion		Fair: Two studies with consistent findings of slightly increased risk for transfusion in TOL although not significant in one.	Moderate
Infection		Poor: Definitions were inconsistent among studies.	Moderate
Incontinence/Pelvic Floor		No studies.	Moderate
Infant Death		Poor: Most studies found increased risk of perinatal death for TOL versus ERCD, but the magnitude of the increase varied greatly.	Least
Neurologic impairment		Poor: Few studies of poor quality.	Least
Respiratory impairment		No studies.	Moderate

Key Question	Study Type*	Quality of Evidence	Suitability of Study Design†
Question 4 What is the incidence of uterine rupture of a cesarean scar, and are there methods for preventing poor clinical outcomes?			
Incidence	II-2	Fair-Poor: Several large cohort studies which were inconsistent in terminology; many with consistent findings of increased risk of symptomatic uterine rupture in TOL versus ERCD.	Moderate
Methods for preventing poor outcomes	II-3	Poor: Few studies, variation in case definition. Fetal bradycardia was frequently associated with uterine rupture; however, inclusion of fetal tracing findings in the definition of uterine rupture makes it difficult to assess the true value.	Least
Question 5 What are the health status and health related quality of life for VBAC and repeat cesarean patients?	None	No studies of women with prior CD.	NA
Question 6 Regarding VBAC and repeat cesarean, what factors influence patient satisfaction/dissatisfaction with their childbirth experience?	III	Fair: Two cross-sectional studies with varied findings.	Least
Question 7 How are economic outcomes related to VBAC, repeat CD, and their respective complications?	Econ	Fair-Good: One good economic model suggests VBAC is cost-effective and provides higher quality of life when chance of vaginal delivery is 76 percent or greater.	Greatest
Question 8 What individual factors influence route of delivery?	II-2	Fair-Poor: Several retrospective cohort studies conducted; all vary in items considered, each with limited adjustment for confounders.	Moderate
Question 9 What factors influence a patient's decisionmaking regarding VBAC or ERCD?	I, II, III	Fair: One good RCT and eight fair quality cohort or cross-sectional studies found women who preferred TOL were more likely to be White, valued the process of labor, and valued social motives such as ease of recovery.	Moderate

Key Question	Study Type*	Quality of Evidence	Suitability of Study Design†
Question 10			
How do legislation, policy, guidelines, provider characteristics, insurance type, and access to care affect health outcomes for VBAC candidates?			
Legislation	II-3	Poor: Few studies that examined only the impact onVBAC rates not safety. None examined the impact of the crisis in malpractice rates on access or safety.	Moderate
Guidelines	I, II	Fair-Good: Several studies with consistent findings that provision of guidelines especially with recommendations of opinion leaders increased VBAC rates; no studies on safety.	Moderate
Provider Characteristics	II	Poor: Several studies, none of which adjusted for differences in baseline risk or potential confounders.	Moderate
Hospital	II	Fair: Consistent findings that teaching hospitals had higher VBAC rates; no comparisons for safety.	Moderate
Insurance	II	Fair: Several studies with conflicting findings.	Moderate

*Study design categories—I: randomized, controlled trials; II-1: controlled trials without randomization; II-2: cohort or case-control; II-3: multiple time series; III: opinions, descriptive epidemiology. U.S. Preventive Services Task Force (1996).

†Suitability of study design categories—Greatest: For comparison studies: Concurrent comparison groups and prospective measurement of exposure and outcome; For rates: population-based or multicenter prospective cohort studies. Moderate: All retrospective designs or multiple pre or post measurements but no concurrent comparison group; Least: Single pre and post measurements and no concurrent comparison group or exposure and outcome measured in a single group at the same point in time. Community Preventive Services Task Force (2000).

Limitations

- Data are insufficient to allow conclusions about the most appropriate delivery choice for a given patient.
- Studies suffered from inconsistent and imprecise definitions for important outcomes.
- Studies frequently failed to ensure comparability between TOL and ERCD groups.
- No study or collection of studies, provide data about the impact of practice variation, provider characteristics, legal considerations such as the effect of rising malpractice rates on the safety of TOL or ERCD.
- The degree to which the association between fetal bradycardia and poor perinatal outcome from uterine rupture rather than confounding by factors detection bias is unclear.
- The degree to which the association between TOL and perinatal death reflects causation rather than confounding by factors such as misclassification of cases, lethal conditions of the fetus, or detection bias is unclear.

Future Research

Future research should focus on conducting methodologically rigorous studies to provide direct evidence regarding the relative benefits and harms of TOL and ERCD. If randomized trials are not done, good-quality studies of TOL versus ERCD must pay attention to the following:

Population. Studies should be conducted in populations of women who are similar in every respect except choice of delivery route (comparability of groups).

Specificity of intervention. Studies should pay close attention to and account for the importance of co-interventions such as use of oxytocin and other medical agents for augmentation or induction of labor.

Precise and standard outcome measures. Variations in reporting of important clinical outcomes were striking.

Studies should consider the following factors in developing outcome measures:

- Etiology. Outcomes such as hysterectomy, infection, maternal mortality, and perinatal mortality must pay specific attention to explicitly identifying the etiology. Lack of precision in this regard allows for both under and overreporting of cases due to misclassification. Examples include whether hysterectomy was performed due to maternal hemorrhage secondary to clinically significant uterine rupture versus hemorrhage due to abruption, uterine rupture through the uterine fundus in a woman with a low transverse incision either due to trauma or other non-incisional causes, and perinatal death due to lethal anomaly versus intolerance or management of labor.
- Standard terminology. In order to accurately measure outcomes, there must be a consistent terminology. Lack of this prevents accurate and meaningful comparisons of risks for each delivery choice. Outcomes such as infection, hemorrhage, and uterine rupture were not consistently defined.
- Separating prevention/prediction strategies from outcomes. As long as potentially important predictors of events such as prolonged fetal bradycardia as a predictor for clinically significant uterine rupture are included in the definition of uterine rupture, their true value as a predictor rather than a confounder will remain unknown.

Predictive Tools

Additional studies are needed to measure the accuracy and yields of existing predictive tools.

Future studies of predictive tools should include measurements of the consequences of false-positive screens and false-negative screens to determine whether there are clinically important harms that result from screening.

Cost

The costs (rather than charges) of labor and delivery and of the surgical processes are poorly understood. Detailed time-in-motion studies would help to estimate these costs.

Availability of the Full Report

The full evidence report from which this summary was taken was prepared for the Agency for Healthcare Research and Quality (AHRQ) by the Oregon Health & Science University Evidence-based Practice Center (EPC), Portland, OR, under Contract No. 290-97-0018. It is expected to be available in the winter 2003. At that time, printed copies may be obtained free of charge from the AHRQ Publications Clearinghouse by calling 800-358-9295. Requesters should ask for Evidence Report/Technology Assessment No. 71, *Vaginal Birth After Cesarean* (VBAC). In addition, Internet users will be able to access the report and this summary online through AHRQ's Web site at www.ahrq.gov.

AHRQ Pub. No. 03-E017
March 2003
ISSN 1530-440X