Objective

To determine if Bishop score is still the best predictor of successful induction of labor.

Background

Pelvic scoring was introduced by EH Bishop in 1964. ¹ In the original article, the recommendation was made to examine women vaginally, starting in the last month of pregnancy, to assess various pelvic factors in an effort to predict onset of spontaneous labor. These factors included cervical dilation, cervical effacement, cervical consistency, cervical position, and the station of the fetal presenting part (Table 1). If the combined assessment was greater than or equal to 7 points, they were recommended to go to spontaneous labor. Therefore, these women’s cervixes were categorized as favorable for induction of labor.

The original research found a 100% success rate of vaginal delivery with score ≥8. Through the years, many studies have looked at weighting components of Bishop score for improved predictability of successful induction.²,³ Some have looked to simplify Bishop score with similar predictive efficacy. Other studies have looked at independent maternal factors such as obesity, gestation diabetes and preeclampsia and their impact on successful induction.⁴,⁵ Throughout all of this research, no scoring system has been shown to be easier and more predictable than the original pelvic scoring developed by Bishop.⁶ As a result, Bishop score is still the standard for induction of labor. It is our goal to address both cervical examination components as well as maternal use and to create an improved scoring system for successful vaginal delivery.

Methods

A retrospective chart review was performed on live, term, singleton, vertex pregnancies where induction of labor was initiated between July 1, 2010 and December 1, 2011 at Kennedy University Hospital. 514 inductions of labor were identified based on delivery logs. 157 patients were excluded for either meeting exclusion criteria (preterm labor, age less than 18, attempting VBAC), incorrectly identified labor induction of labor, or incomplete patient data points, most commonly, incomplete Bishop score. Therefore, 357 patients were included for analysis and their age, race, delivery method, method of induction, degree of laceration, presence of shoulder dystocia, infant weight, 1 and 5 min APGAR scores, gestational age, gravity, parity, indication for cesarean, maternal medical conditions, indication for induction, BMI, and Bishop score data were collected.

The statistical power was computed by setting the α error rate at 5% (2-tailed test) using the N-Query 5.0 computer program (Elashoff JD. nQuery Advisor. Version 5.0 User’s Guide. Los Angeles, CA: Diven Associates, 2002). The power calculations for Bishop score in prediction of successful induction of labor were based upon our clinical observations and/or from the literature. It was assumed cesarean occurs between 14% and 20% after induction (including all methods of induction). The power to detect 2 fold (OR=2.0) effect sizes for outcomes using data from the available charts was calculated (N=400). Univariate statistics were calculated for continuous variables and Chi-square test was used for categorical variables. Student t test was performed to compare each Bishop score to the outcomes (failure or successful induction). Logistic regression analyses were performed when the outcome variable (success or failed induction of labor) is dichotomous, and standardized logistic regression coefficients and their standard errors were used to compute unadjusted odds ratios (OR), adjusted odds ratios (AOR) and 95% confidence intervals (CI). Multiple regression was used to evaluate the relationship between Bishop Score to the outcomes by testing each of components or modifiers such as maternal age, parity, body mass index, fetal weight, presence of meconium, and maternal medical conditions necessitating induction, in an effort to generate an improved scoring system for successful induction of labor. Potential confounders were defined as those factors which alter the adjusted odds ratio by at least 10%. Potential confounding variables including maternal age, ethnicity, parity, cigarette smoking, BMI were tested and controlled for the different models accordingly.

Results

Of the 357 charts reviewed, Bishop scores ranged from 0 to 13, based on Bishop’s proposed method of pelvic scoring from 1964 (Table 1). Bishop score correlated well with the induction of labor success rates, showing a predictable rise from a low of 37.5% at a Bishop score of 0 to 100% with a Bishop score of 13 (p<0.0019). Successful induction rates exceeded 60% with a Bishop score of 3 and 80% with a Bishop score of 7 (Figure 1). With regression analysis, this data projects a calculated success rate equal to 0.0451 x Bishop score + 0.4619. By looking at independent components of Bishop score, successful induction appears to coincide with dilation (p<0.0001), station (p<0.0001) and consistency (p<0.0001) but not amniocentesis (p=0.1109) or position (p=0.5023). However, when adjusting the data for age, BMI and parity, effacement does become statistically significant (p=0.0015). Similarly, as BMI increases, successful induction of labor rate decreases (p=0.0058) (Figure 2).

Parity of one or greater and gravity 2 of two or greater are also more commonly associated with successful induction of labor (p<0.0001 and p=0.0001, respectively). Other analyzed variables, such as age, ethnicity, gestational age and presence of a maternal medical condition were not shown to be reliable predictors of success (p>0.09, p=0.2733, p=0.802 and p=0.4049, respectively).

As a result, a new scoring system was analyzed for comparison to Bishop score. The components of this scoring method included the statistically significant portions of Bishop’s cervical exam (dilation, station and consistency) as well as the other statistically significant variables of gravityparity and BMI (Table 2). When graphing these scores, a plateau is observed at 90%, suggesting a natural cut off of 8 for elective induction of labor and reliable prediction of success (Figure 3).

Conclusions

Though Bishop score is a reliable predictor of successful induction of labor, an improved scoring system for successful induction of labor can be achieved based upon Bishop’s exam of dilation, station, and consistency with additional points given for increasing gravityparity and decreasing BMI. This new scoring method may reduce cesarean rates by improving patient selection criteria for elective inductions. It also may avoid complications of a failed induction of labor by having a lower threshold for cesarean delivery in patients deemed poor candidates by the revised scoring system.

References