OBSTETRICS

Perils of the new labor management guidelines

Wayne R. Cohen, MD; Emanuel A. Friedman, Med ScD

Recent guidelines issued jointly by the American College of Obstetricians and Gynecologists and the Society for Maternal-Fetal Medicine for assessing labor progress differ substantially from those described initially by Friedman, which have guided clinical practice for decades. The guidelines are based on results obtained from new and untested methods of analyzing patterns of cervical dilatation and fetal descent. Before these new guidelines are adopted into clinical practice, the results obtained by these unconventional analytic approaches should be validated and shown to be superior, or at least equivalent, to currently accepted standards. The new guidelines indicate the patterns of labor originally described by Friedman are incorrect and, further, are inapplicable to modern obstetric practice. We contend that the original descriptions of normal and abnormal labor progress, which were based on direct clinical observations, accurately describe progress in dilatation and descent, and that the differences reported more recently are likely attributable to patient selection and the potential inaccuracy of very high-order polynomial curve-fitting methods. The clinical evaluation of labor is a process of serially estimating the likelihood of a safe vaginal delivery. Because many factors contribute to that likelihood, such as cranial molding, head position and attitude, and the bony architecture and capacity of the pelvis, graphic labor patterns should never be used in isolation. The new guidelines are based heavily on unvalidated notions of labor progress and ignore clinical parameters that should remain cornerstones of intrapartum decision-making.

Key words: active phase, arrest of descent, arrest of dilatation, arrest of labor, deceleration phase, dysfunctional labor, labor curve, partogram

The seemingly inexorable increase in the use of cesarean delivery, and the substantial contribution that dystocia and related diagnoses have made to that increase, have prompted a reevaluation of what constitutes normal labor.[1-4] As a result, new guidelines promulgated jointly by the American College of Obstetricians and Gynecologists (ACOG) and the Society for Maternal-Fetal Medicine (SMFM) were released.[5] The new recommendations define abnormal labor and provide guidelines for its management that differ sharply from those originally described by Friedman,[6-10] which have formed the basis of the clinical management of labor for many decades in the United States and elsewhere. For that reason, a thorough analysis of the proposed standards is warranted to ensure that changes recommended for obstetric care during labor are justified by the available evidence.

The guidelines are based heavily on analytic methods used by Zhang and colleagues[11-14] to describe the patterns of cervical dilatation and fetal descent as functions of time elapsed in labor. Their findings, which have been rapidly adopted in some parts the obstetric community, have not yet been validated. For the reasons we briefly summarize

in this commentary, we believe the new ACOG/SMFM recommendations provide definitions of dysfunctional labor and guidelines for its management that, however well intentioned, are likely to impose undue risk on mother and fetus.

Historical background

Prior to the mid-1950s, the evaluation of progress in labor was based primarily on its duration. Vague admonitions such as, "Never let the sun set twice on a laboring woman," which were based on prevailing observations about average labor duration and outcomes,[15] were commonly intoned. This approach was, however, ineffective in identifying when intervention would be appropriate or optimal.

In 1954, the first of hundreds of studies of labor by, or based on the work of, Emanuel Friedman[6] was published. Friedman's work built upon previous investigators' attempts to describe the events of labor as a function of time.[16-18] Their recognition of the practical implications of this approach was hampered by what we now know to have been erroneous assumptions about labor, particularly with regard to the role of membrane rupture. The first publications[6-8] describing the graphic patterns of dilatation and descent stimulated the interest of many investigators, and led to the formulation of criteria that made the assessment of progress in labor objective rather than arbitrary.[9,10,19-26] Unfortunately, the criteria have not always been applied appropriately, in part because of some misunderstandings about the curves and their proper place in clinical care.

Misconceptions

It has often been alleged that Friedman's seminal observations regarding the labor curves rest on a fragile foundation because they were never corroborated by others. In fact, numerous studies done

From the Department of Obstetrics and Gynecology, University of Arizona College of Medicine, Tucson, AZ (Dr Cohen); and the Department of Obstetrics, Gynecology, and Reproductive Biology, Harvard Medical School, Boston, MA (Dr Friedman).

in different parts of the world over the course of several decades confirmed the basic nature of the original curves, and validated their usefulness in clinical practice.[27-43] There have been disagreements over the importance of the latent phase or even the existence of the deceleration phase of dilatation, but the core finding that active-phase cervical dilatation progresses linearly, with a lower limit of normal approximately 1.0 cm/h in nulliparas, has been remarkably consistent among studies. It is also noteworthy that in many institutions the introduction of labor curves to clinical care was associated with a decline in the cesarean rate.[30,33,34]

Some of the early data were collected using a mechanical cervimeter to obviate the potential subjectivity in clinical examination,[9,44,45] and cervimetry by investigators using various tools confirmed the sigmoid nature of the dilatation curve.[46-49] Limited data from more recently developed techniques to automate cervical assessment also appear consistent with the earlier observations.[50,51] Sigmoid-shaped curves of cervical dilatation have even been described in cows, suggesting a common pattern of labor among mammalian species.[52]

Given the large body of evidence confirming the basic pattern of progress in normal labor, it is difficult to believe that labor progresses very differently today from how it was originally described. Why, then, do the labor curves of Zhang and his colleagues differ from those of previous observers? One explanation was provided by Zhang himself when he and his colleagues applied their analytical methods to the very same data Friedman had analyzed from the Collaborative Perinatal Project.[14] Friedman's analysis of those data revealed a sigmoid-shaped dilatation curve; that of Zhang et al revealed an exponential curve, essentially the same as they had found from contemporary labors. Clearly, what had changed was not the nature of progress in labor, but how the data were analyzed. This raises the question of which analytic technique provides a more accurate model of labor progress: that of Friedman or that of Zhang et al?

In trying to address that question it is important to understand that the original dilatation and descent curves were based on and confirmed by direct experimental observations made on women in labor. The primacy of direct observation over theoretical conceptualization or indirect analysis of data in hypothesis testing has been a central tenet of the scientific method since the Enlightenment. When the results of an analytic approach differ from those derived from observation, it is important to understand why this has occurred, and try to adjudicate accordingly, before declaring the direct objective findings invalid.

Analytical issues

The labor curves in Friedman's original reports were not created by using complex mathematical formulae, as some have suggested.[2] The initial data were collected by a single observer.[6] Subsequently, data from multiple practitioners in a single institution were reported.[7,8] In both instances, the curves were drawn by hand, the descriptions were empiric, and the statistical analysis basic. Only later was a more sophisticated method of assessing the labor graphs by computer used to analyze >10,000 nulliparas from multiple institutions.[53-56] This more sophisticated analysis confirmed the initial findings regarding the nature of the cervical dilatation and head descent time functions.

The computer algorithm used was developed with the Office of Biometry of the National Institutes of Health. Raw labor data were plotted on a probit (ie, the normal probability) scale, to convert the sigmoid curves to straight lines.[57,58] The maximum slope data were converted to logarithms to normalize their right-skewed distribution. The linearity thus achieved made the data amenable to descriptive statistical study for determining distributions and limits of normal, which have until recently stood the tests of time and clinical applicability.

By contrast, Zhang and colleagues used a high-order polynomial curve-fitting program to analyze dilatation and descent data, and interval-censored

regression to fit curves based on centimeter-by-centimeter median traverse times. We have concerns about the application of this technique to labor.

We do not profess personal expertise in this area, but we are impressed by the negative comments and strong skepticism encountered in the engineering literature pertaining to the limitations of high-order curve-fitting methods.[59] Such models do not guarantee reliable results. Indeed, high-order curve fitting may not be appropriate or even necessary for most situations. Low-order quadratic curve fitting is preferable, whenever possible, and yields results that are at least as accurate. In fact, the higher the order, the less satisfactory curve-fitting accuracy tends to be. This is so because 'noise' (ie, unstable data points, especially if those points are spread apart from each other or are located at the ends of the range of data) is magnified. As a consequence, portions of the derived curve are distorted. In this regard a leading authority opined that, "It is important to keep the order of the model as low as possible…As a general rule the use of high-order polynomials (k >2) should be avoided unless they can be justified for reasons outside the data…Arbitrary fitting of high-order polynomials is a serious abuse of regression analysis."[59] Zhang et al used polynomial curve-fitting models of the order of 8-10, far in excess of the cited recommendation of no higher than 1 or 2.

Other investigators have used interval data to create labor curves, with varying results. Gurewitsch et al[60] found a sigmoid curve of dilatation, but Chen and Chu[61] found results similar to those of Zhang et al in terms of curve shape and much lower rates of dilatation.

Thus, the differences alleged to exist between the Friedman and the Zhang curves are likely due to the different mathematical models used to fit these curves. This is confirmed by Zhang's finding, noted above, that the same data Friedman and Neff[62] analyzed decades ago yielded exponential curves with the curve-fitting methods used by Zhang and his colleagues.[14]

The approach by Zhang et al is likely to have introduced an important set of