



99.5%

False Positive Rate?

Absurd!

By Max Freeman & Rachel Gross

absurd: utterly or obviously senseless, illogical, contrary to all reason or common sense, laughably foolish or false. (In *Merriam-Webster.com*, retrieved June 26, 2015, from <http://www.merriam-webster.com/dictionary/absurd>.)

Defense lawyers have a propensity for using “sound bites” from current literature to try to excuse their clients’ carelessness. One recurrent example is using the claim from Clark and Hankins about false positive rates for electronic fetal monitoring to try to explain away findings on a fetal monitoring strip. They claim: **A test leading to an unnecessary major abdominal operation in more than 99.5% of the cases should be regarded by the medical community as absurd at best.**¹ Ironically, by this logic, a caesarean section in which the fetus was severely acidotic and depressed at birth and went on to develop CP would be a “true positive”. It seems clear then that physicians motivated by the desire to *eliminate all true positives* are not waiting long enough to satisfy the criteria for morbidity and avoid the stigma of a false positive.

Clark and Hankins go on to state:

¹ Clark, Steven L. & Hankins, Gary D. V. (March 2003). “Temporal and Demographic Trends in Cerebral Palsy – Fact and Fiction.” *Am. J. Obstet Gynecol*, Vol. 188, No. 3, pp. 628-633.

Of course, the above conclusions regarding the inefficiency of electronic fetal monitoring in the prevention of cerebral palsy are predicated on the assumption that most obstetricians are well-trained and competent and that caesarean sections based on abnormal fetal heart rate tracings are, for the most part, performed in a timely manner in accordance with the traditional teachings regarding monitor interpretation. We feel an assumption of general competence among board-certified obstetricians–gynecologists is justified.²

² Id.. We could substitute for the word “competent” the word “bewildered”. What is one to make of the pronouncements of the ACOG regarding EFM? The various permutations in the classification of FHR patterns, the continuing modification of intervention guidelines, without a single mention of the notion of preventable injury provides no guidance (nor is it likely intended to). There is considerable reason to believe that most obstetricians are competent but lacking in support and

Without commenting on – and certainly without accepting – the second assumption, the first deserves some analysis. A statement so bold compels one to wonder: where does this assumption come from?, is the assumption valid? what does it mean?, and does it have any practical significance?

It appears that the notion of an *extremely* high false positive rate in association with fetal monitoring first came from a paper published in March 1996 in the New England Journal of Medicine by Karin B. Nelson and others, entitled “Uncertain Value of Electronic Fetal Monitoring in Predicting Cerebral Palsy.”³ The article is interesting in several respects. First, it should be noted that neither Dr. Nelson nor her co-authors appear to have any expertise in actually reading electronic fetal monitoring strips. Dr. Nelson is a pediatric neurologist. Her co-authors are not medical doctors at all. In any event, on closer analysis of the article, that fact may well not be particularly significant. As stated in the article: “NO MONITORING STRIPS WERE AVAILABLE FOR REVIEW.”⁴

If no fetal monitoring was available for review, one would anticipate that at least the authors had reviewed the observations made by the nurses who had a duty to periodically review and record findings on the monitor. Instead, alas, “the findings on fetal monitoring that were recorded were those noted in the birth records by the physicians attending the deliveries.”⁵ Dr. Nelson, simply assuming that from

education from the “vague” guidelines offered by ACOG (Clark). Consider also the following comments in a recent article on the management of Category II patterns authored by 18 luminaries in the field: “Unfortunately, this body of work [EFM research] has primarily served to raise more questions than it has answered.” “As a medical community, we seem to know less than we thought we did 30 years ago regarding the utility of this ubiquitous technique [EFM].” Clark, et al., “Intrapartum Management of Category II fetal heart tracing towards standardization of care”, *Am. J. Obstet Gynecol* 2013.

³ Nelson, Karin B., et al. (March 1996). “Uncertain Value of Electronic Fetal Monitoring in Predicting Cerebral Palsy.” *N. Engl. J. Med.*, Vol. 334, pp. 613-619.

⁴ Id. See also Schifrin, Barry S., et al. (June 1996). “Correspondence: Letters to the Editor in Response to “Electronic Fetal Monitoring in Predicting Cerebral Palsy”.” *N. Engl. J. Med.*, Vol. 335, No. 4, pp 287.

⁵ Id.

1983 through 1985 the physicians in the San Francisco Bay Area actually had a practice of periodically reviewing and recording in the chart findings on the fetal monitor, then compared those findings to children in the Bay Area with cerebral palsy. “Cerebral palsy was defined as a chronic disability originating in the central nervous system, characterized by aberrant control of movement or posture, appearing early in life and not resulting from progressive disease. ... Children with mild involvement or isolated hypotonia [whatever that means] were not included.”⁶ The children’s demographic and clinical data were obtained from birth certificates and medical records at more than 40 hospitals. The data were abstracted by nurses working at the California Birth Defects Monitoring Program who “did not know that the study was about cerebral palsy.”⁷ From even this limited analysis, Dr. Nelson nevertheless found that “multiple late decelerations were associated with nearly a quadrupling of the risk of cerebral palsy ... And decreased beat to beat variability with nearly a tripling of the risk. ... The occurrence of multiple late decelerations, decreased beat to beat variability or both abnormalities was associated with a sharp increase in the risk of cerebral palsy.”⁸ But late decelerations and alterations of variability may occur in many other clinical situations unrelated to any significant threat of neurological injury. Late decelerations, even those accompanied by alterations in variability, are commonplace with maternal supine hypotension, excessive uterine activity, or induction of epidural anesthesia. Most often, however, those aberrations are readily correctable by conservative maneuvers.

The findings by Nelson, et al., regarding what might go wrong on a strip are potentially important. Maybe, they are even important enough to actually merit looking at the fetal monitor strips. Dr. Nelson of course took a different approach. Rather than focus on the actual clinical findings and the potential for recovery, Dr. Nelson instead went on to extrapolate her findings to the entire population of children born during those years studied in all the countries observed. From that, she multiplied out her findings and concluded “the estimated false positive rate is 99.9% among children with none of the other

⁶ Id.

⁷ Id.

⁸ Id.

risk factors we examined and 99.6% in the high risk group.”

It should be noted that this is not Dr. Nelson's first excursion into statistical fantasy-land as it relates to birth-injured children. In her 1986 article “Antecedents of Cerebral Palsy: Multivariate Analysis of Risk,” Nelson and her co-author, Ellenberg, examined approximately 54,000 pregnancies between 1959 and 1966 at 12 separate university hospitals. These children were examined until they reached seven years of age. Only 189 of these children had cerebral palsy. By using logistical regression, Nelson and Ellenberg purportedly identified 25 leading risk factors among these 189 cerebral palsy cases, compared this statistical ratio to the supposed equivalent of a larger general population of the same study, and then divided these factors into two temporal stages. Nelson and Ellenberg used factors for Group or Stage One that were preexisting by the time labor began, including motor deficit in an older sibling, maternal seizures and/or bleeding, and fetal birth weight. The second group consisted of risk factors that began after the onset of labor, including neonatal seizures and infections, low fetal heart rate in labor, and time elapsed before the first cry. (We set aside for this discussion the problems with the validity of her underlying assumptions about risk factors in any medical sort of way.) She found that Group One accounted for 34% of the 189 cerebral palsy cases and Group Two accounted for 37% of those same cases. Based upon this small discrepancy, Nelson made the leap to simply conclude that with respect to the role of any labor and delivery factor to account for cerebral palsy, “when characteristics were considered in the sequence in which they arose, no factor in labor or delivery was a major predictor.”⁹

This was later examined in “The Temporal Stage Fallacy: a Novel Statistical Fallacy in the Medical Literature,” in which Shier and Tilson pointed out that Nelson’s 1986 article is entirely fallacious. Shier quite accurately points out that Nelson’s study automatically gives deference to early factors like genetics over later occurring ones like fetal distress. Even if we were to ignore the fact that Nelson simply glossed over her contextual definition of “substantial” difference between the two groups, Shier points out

⁹ Nelson, K and J. Ellenberg (1986). “Antecedents of Cerebral Palsy: Multivariate Analysis of Risk.”. *N. Engl. J. Med.* Vol. 315(2), pp. 81-86.

that the logic is inaccurate because it is based only a small subset population. In order to be accurate, Nelson and Ellenberg would have had to state that the study “did not identify a substantially larger number of *new* cases.”¹⁰ The ultimate putdown of this argument came from Paneth, who showed that using the same principle in pirate executions; you could conclude that those executed by drowning actually died because of “walking the plank”.¹¹ One could argue that the strong correlation of fhr patterns and subsequent CP – “true positives” – represented failures of obstetrical intervention, an issue about which Nelson is conspicuously silent. This feckless study by Nelson did, of course, gain traction within the ACOG community when it published its 2003 pamphlet claiming that complications during childbirth like asphyxia almost never cause cerebral palsy. Not so shockingly, Nelson and Ellenberg’s 1986 study was cited in support.

With respect to the claim that electronic fetal monitoring has an almost 100% false positive rate to predict cerebral palsy, one must step back and think about what is being examined to begin with. Fundamentally, it would seem that there is some sort of difference between medical “monitoring” and a medical “test”. Monitoring is defined as: “to observe and check the progress or quality of something over a period of time; to keep under systematic review; to keep under observation.”¹² A test is defined as: “to take measures to check the quality, performance, or reliability of something, especially before putting it into widespread use or practice; to take measures to determine result.”¹³ One would think that if, indeed, conceptually one was looking for a false positive or a false negative rate **for a test**, the subject of the study would be something that actually **tests** for the condition about which statistics are being derived. In determining the value of a “test”, it is important to note the prevalence of the disorder you are testing for and NOT attempt to change the outcome.

¹⁰ Shier, David and J. Lee Tilson (2006). “The Temporal Stage Fallacy: A Novel Statistical Fallacy in the Medical Literature”. *Med. Health Care Philos.*, Vol. 9, pp. 243-247.

¹¹ Paneth, N., “Birth and the origins of cerebral palsy”. *N. Engl. J. Med.* 1986; 315: 124-6.

¹² “monitor” In *Merriam-Webster.com*. Retrieved June 26, 2015, from <http://www.merriam-webster.com/dictionary/monitor>.

¹³ “test” In *Merriam-Webster.com*. Retrieved June 26, 2015, from <http://www.merriam-webster.com/dictionary/test>.

Even Dr. Nelson could not go so far as to suggest that an electronic fetal monitor was a cerebral palsy *test*. Not only would that be absurd under any definition, it is simply not the purpose for which the monitor was designed. An electronic fetal monitor merely observes the fetal heart rate and the parameters about it which can be measured. These include baseline, variability, accelerations, decelerations, and evolution of pattern designed to inform about fetal asphyxia. The monitor also observes the patterns of maternal contractions. If one is going to question the false positive or false negative rate of such a monitor, it would seem the inquiry would be about how accurately the monitor examines the parameters for which it was designed. Assuming that the monitor *does* reasonably measure what it was designed to measure, the question really becomes a matter of interpretation and the timing of intervention, not false positives.

Nelson's standby argument revolves around the concept of unnecessary caesarean sections. Even if we were to presume that a caesarean birth is the universal response to fetal distress shown on a fetal monitor, Nelson's arguments imply that if a child is delivered via caesarean and does not have cerebral palsy, then that caesarean section is totally without benefit. She never seems to consider that the very caesarean birth she so disdains may have actually saved the infant's life and *that* is the reason, in fact, why the infant does not have cerebral palsy. Think only of the study that would have to be designed to test Dr. Nelson's hypothesis. The lack of cerebral palsy could simply mean that those doctors had avoided disaster. Perhaps a better understanding of fetal heart rate patterns and more uniform and discriminatory analysis of these same patterns than prevails today is more innovative rather than focusing on the manner of birth itself.¹⁴

Another way of looking at this quirk in Dr. Nelson's reasoning is to focus on her unstated assumption that there will be no intervention depending on the finding on the monitor. The corollary assumption is, if there are abnormal findings on a fetal heart monitor, a caesarean section is performed and the child does not have cerebral palsy, then *a priori*, the caesarean section was unnecessary. By analogy, in most surgical suites employing anesthesia, a patient's blood pressure is monitored

intraoperatively. One reason that the blood pressure is monitored is that if indeed the patient's pressure drops precipitously during the procedure, (either because of hemorrhage, reaction to anesthetic agent, etc.), and if it is not corrected the patient could suffer unnecessary injury or die. In most operative procedures in the country today, if a patient's blood pressure drops during surgery, the matter is addressed and treated. Statistically, Dr. Nelson would say that blood pressure monitoring has an incredibly high false positive rate for injury or death. Accordingly, by this chop-logic, one could question the rationale of monitoring blood pressure in surgery today.

Even if Dr. Nelson's analysis could be presumed to have any statistical validity whatsoever, consider the use of its conclusion. By way of analogy, one would be hard-pressed to suggest that teenage drunk drivers are not a serious threat to themselves and others on the road. Would anyone seriously suggest that when a collision occurs involving a drunk teenage driver, that the intoxication did not play some role? Still, statistically, the false positive rate for teenage drunk drivers in this country on any given weekend is significantly higher than the false positive rate which Dr. Nelson attributes to fetal monitoring. We don't suggest for a moment that Dr. Nelson, Dr. Clark, or Dr. Hankins would roll those dice and encourage their daughters to ride around with a drunk teenage driver on a Saturday night. Rather than downplaying the importance of fetal monitoring, more emphasis should be placed on improving the interpretation of the monitoring parameters and improving the understanding of basic maternal fetal physiology in an effort to actually keep patients out of harm's way.

**Max Freeman
& Rachel Gross**
Miller Weisbrod, LLP
11551 Forest Central Dr
Suite 300
Dallas, TX 75243

¹⁴ Schifrin, *supra*.