MEASURING BIRTH OUTCOMES: VALIDATING THE PERINATAL OUTCOME INDEX

ARTICLE

Patricia McNiven, RM, BScN, Msc, PhD
Karyn Kaufman, RM, BSN, MS, DrPH
Murray Enkin, MD, FRCSC, FACOG, LLD, Dsc

ABSTRACT

We undertook a study to assess the reliability and validity of a new measure of low risk maternity care. A Perinatal Outcome Index (POI), which combines intrapartum process of care and clinical outcome items into a summary index score, was originally developed and evaluated in the Netherlands. It was designed to measure the extent to which a labour and birth are "optimal", that is, one with minimal intervention resulting in a healthy mother and a healthy baby. We modified the Dutch index to make it applicable to a Canadian setting. A panel of experts who were not connected with the study reviewed the modified version for applicability, feasibility of obtaining data easily, and content validity. Data were abstracted from the health records of 324 women in one hospital and two midwifery practices to obtain Perinatal Outcome Index scores and examine aspects of construct validity. We measured the inter-rater reliability of the research assistants who abstracted information. The panel achieved consensus on all items in the modified Perinatal Outcome Index to establish content (face) validity. Labour and birth data were readily obtained from health records with high inter-rater reliability (Kappa 0.78). In a linear regression model, birth at home, multiparity, and having a midwife or family physician as a care provider were significantly associated with higher scores (having a more optimal birth) and accounted for 37% of the score variance. The Perinatal Outcome Index has satisfied our expectations for content and construct validity. Research assistants found it easy to use and data items were readily available from women's health records. Inter-rater reliability was acceptable. We believe the modified index will be useful for comparative studies among women at low or average risk, and for quality assurance programs.

KEY WORDS

perinatal outcome, outcome measures, low risk obstetrics, midwifery evaluation

THIS ARTICLE HAS BEEN PEER-REVIEWED.

BACKGROUND

New approaches to maternity care require careful comparison and review, but we still do not know how to best evaluate them. Traditional outcome measures, such as perinatal mortality or serious morbidity are, thankfully, too rare to be useful in studies of women at low or average risk. More attention is now paid to the quality of the birth experience, but this is difficult to measure. Maternal satisfaction is important, but is only one component of the experience.

What is required is a comprehensive, sensitive index that integrates both process and outcome variables into a readily usable measure of maternity care, one that can discriminate not only between the inadequate and the good, but also between the good and the best. We searched for such an index to use as an outcome measure for a study to compare midwifery with physician led maternity care. We wished to use the optimality index for a cost-effectiveness study of midwifery care, because no such study has been reported using validated outcome measures. A cost effectiveness study must have a comparison group to be meaningful. We chose physician care as our comparison group. The index that most closely met our criteria was a Perinatal Outcome Index (POI), developed and evaluated in the Netherlands.[1]

Scoring systems such as the Apgar score or the Bishop score are commonly used in obstetrics as part of clinical care. Scoring systems are also used to summarize clinical information for research purposes. In rehabilitative medicine, quality of life assessment scales are used in clinical trials to measure the effects of various interventions because a single outcome measure is not sufficient. Research about the effects of care on low risk pregnancies requires an important primary outcome that occurs with enough frequency to be observed in reasonably sized study groups. The purpose of the Perinatal Outcome Index is similar to the quality of life assessment in that it encompasses many factors because no important outcome occurs often enough to be a good research outcome.

CONCEPTUAL MODEL

The Dutch Perinatal Outcome Index was designed to measure "maximum outcome with minimal intervention".[1] The index is based on the "optimality concept," first described by Prechtl in 1968.[2,3] It is a clinimetric index, designed to combine multiple variables into a single outcome expression that rates complex clinical phenomena.[4] A clinimetric index is a health measurement scale that combines a number of discrete factors into a summary score. In studies of interventions that focus only on outcomes such as serious morbidity and perinatal mortality, measurement is straightforward. However, there are factors in clinical research that make the situation more complex.[5] Collecting factors relating to process, in addition to absolute outcomes, has become more important. Additionally, in Canada, serious morbidity and mortality are infrequent. A study will require a very large sample size when these factors are used as the primary outcome. This would be difficult in a study of Canadian midwifery. The major innovation of the Dutch index is the manner in which it combines process and outcome variables. Rather than trying to define different degrees of pathology in medical terms, Weigers et al chose to describe an "optimal birth".[1] This was defined as a birth that occurs at the end of a full term pregnancy, and results, without any interventions or complications, in a healthy baby and healthy mother. →

Most women and their caregivers hope for such a birth: one with no complications, no interventions and no risk factors. This "optimal birth" is the basis for the index described in thisstudy. The Perinatal Outcome Index does not evaluate the appropriateness of the interventions that women may receive. However, because the use of an intervention signals that a clinical indication arose during the labour, scores achieved on the Perinatal Outcome Index will be lower when interventions are used. Any intervention in labour is indicative of a deviation measure quality of care; it measures the optimality of the birth from optimal and is scored accordingly. The index does not experience.

An optimal birth, as defined in the Perinatal Outcome Index, receives a maximum score of 40 points. One point is taken away for each item that does not meet the description of "optimal". The sum of the remaining points is the final score. A higher score means that the birth is closer to the optimal situation (no complications, no interventions and no risk factors). The individual components of the index are not weighted; the index is self-weighting because many items are correlated and thus points are lost simultaneously.[1,4] For example, if a labour had slow progress, a point is removed for length of labour. If augmentation and amniotomy were used to aid progress, then a point is removed for each of these items. If an intravenous infusion and epidural analgesia are initiated, then a point is removed for each these items as well. Thus, for this example, the score would be reduced by at least five points from the maximum value.

The combined process and outcome score is useful in research studies because many individual outcomes are too infrequent among low risk populations to permit statistical significance to be achieved with a reasonable sample size (since sample size is based on the frequency with which an event of interest occurs). The Perinatal Outcome Index combines individual items into a composite score. This then becomes the primary outcome for research studies and permits relevant research to be carried out with a feasible sample size.

It is important to remember that the index is designed for research, rather than clinical use. It does not assess the appropriateness of the care given to an individual woman, nor does it assess her personal experience. Rather, it is designed to compare maternity care services among different groups, or at different sites. The index score is not designed to replace analysis of serious adverse outcomes, which should always be itemized, analysed and reported.

We could not assume that the Perinatal Outcome Index as developed in the Netherlands would apply to a Canadian population. Wiegers et al state that, "[the index] should not be regarded as a static instrument and, if necessary, be adapted to changed insights in obstetric and maternity care."[1] We decided to modify the index, to adapt it for an evaluation of maternity outcomes in a Canadian setting, and to evaluate the applicability, reliability and validity of the modified index, before using it for our proposed comparative study. The current study was designed to test the properties of the modified Perinatal Outcome Index.

RESEARCH QUESTIONS

1. Applicability and feasibility: Are items in the index applicable to the Canadian situation? Can the information be obtained easily from the health records of women?

2. Content Validity: Is there evidence to support changing item descriptors or deleting an item from the Index?

3. Construct Validity: Is there a difference in scores between groups of women whom we expect to have more or less optimal births?

4. Reliability: What is the inter-rater reliability of the scores?

METHODS

We consulted with members of the research team and an invited panel of experts to address the questions of applicability, feasibility and content validity. Research assistants abstracted data from selected women's health records in one participating hospital and two participating midwifery practices. The Kappa coefficient was used to evaluate inter-rater reliability. Each aspect of the study is reported below with a more complete description of its methods and results.

CONTENT VALIDATION

We reviewed and modified the individual items of the index for use with a Canadian population. Content or face validity was established by having the modified index reviewed by a panel of experts. The panel included a consultant obstetrician, a midwife, a family physician, a nurse, and a consumer, none of whom were members of the research team. The consumer was the leader of a maternity care advocacy group.

The index items, with their definitions and some background materials, were circulated to the panel. A rating form, indicating expert agreement or disagreement and the appropriateness of the definition of each item was sent with the materials. The panel then met to review the items and reach consensus on the inclusion of each item. The panel of experts and the research team agreed upon the definitions for each item.

The modified index consists of 40 items arranged into three sub-scales (Table 1). Each item has a description of its optimal state, for example, the description of an optimal Apgar score is greater than or equal to nine at five minutes. Definitions or descriptors suitable for Canadian practices were provided for all items. We modified and combined several items from the Dutch index and deleted others. For example, mastitis, cystitis & endometritis were combined to form one item, postpartum infection. Referral during labour was deleted because obstetricans may provide primary care and would not refer in situations where family physicians and midwives would be →

required to do so. The panel also added several items to the index. Intrapartum antibiotics, epidural analgesia and intravenous infusions were included because the use of these interventions indicates the presence of a complication or risk that detracts from the optimal birth. The panel judged that some items would be infrequently applicable in the local setting (pudendal block, routine enema and pubic shave) but because these interventions continue to be used in some Canadian locations, the items were retained. When the panel reached consensus on all the changes, the research team then judged the modified Perinatal Outcome Index to have face validity for a Canadian population. We reviewed the items to determine whether they would be retrievable from hospital or home birth records and determined they would be part of the usual documentation.

The original Dutch Perinatal Outcome Index measured background and demographic factors, but we decided not to include these because they describe the study population and do not relate to intrapartum care. The modified index is an intrapartum and early postpartum index. We did not include items referring to pregnancy complications or prenatal care because we felt these items should not be combined with intrapartum items into one large index. Furthermore, in our setting, antenatal data cannot always be obtained from hospital records.

STUDY SAMPLE

To test construct validity we abstracted chart information from a convenience sample in an urban tertiary care hospital. We selected charts of women who would have been considered eligible for midwifery care because we wished to test the Perinatal Outcome Index among a group who were not at high obstetrical risk. We did not sample from a high risk population because the index was not designed for high risk situations where adverse outcomes are more frequent. We did not attempt to assemble a sample that was representative of the overall population of women giving birth in this community, as we were testing the index, not the population. The chart information was used to calculate a score for each subject and to provide data about other variables of interest. We estimated that a sample size of 200 would be sufficient to test the properties of the index. Since this was an initial assessment a power analysis could not be performed. We over-sampled among the eligible charts to compensate for potential missing records. We selected charts from eligible women who gave birth between August 1998 and May 1999. All women admitted by a family physician or midwife during the time period were selected for the sample. The next eligible woman admitted by an obstetrician was then selected. Women were retained in their admission group even if their care was eventually transferred to an obstetrician, for example, for Caesarean section. No information regarding the labour or birth was considered during subject selection. Overall, 283 records were available→

TABLE 1:
CANADIAN PERINATAL OUTCOME INDEX

Intrapartum

1 no analgesic medication during first stage
2 no epidural
3 time between rupture of membranes and birth ? 18 hours
4 no intrapartum antibiotics
5 no meconium in the amniotic fluid
6 first stage of labour ?12 hours
7 second stage not prolonged
8 vertex presentation at birth
9 no perineal laceration
10 no episiotomy
11 no dystocia
12 reassuring fetal heart rate pattern
13 no induction of labour
14 no augmentation of labour
15 no instrumental delivery
16 no pudendal block
17 no Caesarean section
18 no manual removal of placenta
19 no excessive blood loss during delivery
20 no blood transfusion
21 no venipunctures
22 no intravenous infusion
23 no continuous electronic fetal monitoring
24 no routine shave or enema
25 no other problems or interventions

Newborn

26 birth weight between 10th and 90th percentile
27 Apgar score ?9 at 5 minutes
28 no perinatal death
29 no admission to neonatal unit
30 duration of gestation ?37 weeks
31 no significant congenital anomalies
32 no birth trauma
33 no glucose or formula in first 48 hours
34 initiated breast feeding
35 no separation of mother and baby
36 no other problems in the first 48 hours

Postpartum

37 no infection
38 no medication in puerperium
39 no bladder dysfunction
40 no other maternal postpartum problems

for abstraction. The final sample included 56 women cared for by midwives, 63 by family physicians and 164 by obstetricians. Of the latter group, 53 women received care from a model of shared antenatal care between family physicians and obstetricians. We sought an additional convenience sample of records from the midwifery practices in the city for women who had a home birth in the same time period as the hospital births. Records for 41 home births were available for abstraction.

PROFILE OF STUDY SUBJECTS

We abstracted information on socioeconomic status, level of education, pregnancy history and baseline health. These demographic data were not always available from the hospital records. Descriptive information about the study sample is shown in Table 2. More than 83% of women in the sample had an identified partner. Nearly 40% of those whose educational level was charted had post-secondary education, but this information was often not noted on obstetrical records. Almost 90% of women were in the 21-34 year age group. Nearly two-thirds of the sample were multiparous women. There were no intrapartum or early neonatal deaths.

TABLE 2:
BACKGROUND DATA (n=324)

Variable	*Frequency n(%)*
Parity	
Primipara	96 (29.6)
Multipara	207 (63.9)
Missing data	21 (6.5)
Support	
Identified partner	(271 (83.6)
No identified partner	23 (7.1)
Missing data	30 (9.3)
Educational Level	
Some high school	12 (6.5)
Completed high school	32 (9.9)
Post-secondary education	126 (38.8)
Missing data	154 (44.8)
Maternal Age	
under 20 years	20 (6.2)
21 - 34 years	267 (88.4)
over 35 years	32 (9.9)
Missing data	5 (1.5)
Smoking	
Nonsmoker	238 (73.5)
Smoker	56 (17.3)
Missing data	30 (9.2)

DISTRIBUTION OF PERINATAL OUTCOME INDEX SCORES

Index scores were calculated from the 324 charts that were reviewed. The scores ranged from 16 to 39 out of a possible 40 points. The mean score was 30.3, the median was 31 and the mode was 36. The scores were not normally distributed, but skewed toward the higher end of the index. This was expected since most births are close to optimal. There was, however, reasonable variation in the scores among this study sample.

ASSESSMENT OF CONSTRUCT VALIDITY

We wanted to examine the association of index scores with variables that are known to influence perinatal outcomes. We thought that place of birth, parity, type of caregiver, and some demographic variables would influence the score. We expected women who had a home birth to have a higher score than women who gave birth in the hospital, and multiparous women to have higher scores than nulliparous women. We expected that the scores for women cared for by midwives and family physicians would be higher than for those cared for by specialist obstetricians. Similarly, we expected that women with a primary support person and women who were nonsmokers would have higher scores.

As shown in Table 3, the mean index scores varied significantly with site of birth, parity and caregiver. Women who had a home birth scored higher than those who gave birth in hospital. Multiparous women scored higher than nulliparous women. Women cared for by family physicians or midwives had a significantly higher mean Perinatal Outcome Index score than those receiving care from obstetricians. Nonsmokers and women with a partner had higher mean scores than smokers or women with no identified partner, but the differences did not reach statistical significance.

We performed a multiple stepwise regression analysis in order to determine the relative importance of the variables that were associated with the Perinatal Outcome Index scores. As shown in Table 3, the full set of variables entered into the equation accounted for 37% of the total variance. →

continued from page 10...

(The R^2 value indicates the amount of variance in the index score accounted for by the individual variable, holding the other effects constant. A higher R square value indicates a stronger association.)[6] Nearly 30% of the variance was accounted for by place of birth. Caregiver and parity accounted for another small proportion. The other variables, smoker and personal relationship, did not significantly explain score variance. A large proportion of the total variance was not accounted for in the statistical model. This was not surprising because of the many factors that can influence birth outcomes.

Because site of birth accounted for most of the explained variance, and because it is likely to be confounded with type of caregiver and parity, we ran a separate analysis restricted to hospital births, to determine whether these latter two variables would remain significantly associated with the index score when site was a constant. Caregiver and parity remained as significant variables in the regression model. Maternal smoking and support did not account for a significant increase in the R^2 value.

We sought to test criterion validity by a concurrent comparison of the Perinatal Outcome Index with some other validated index. We could not, however, find another measure that could serve as a "gold standard". Thus, we relied on face and construct validity as indicators that the index actually measures what is intended by the concept of optimality.

RELIABILITY

Perinatal Outcome Index scores must be reproducible and consistent if they are to be used as outcome measures. We measured the inter-rater reliability of the scores. Research assistants were trained to abstract data from the women's health records by an investigator who was familiar with the organization and terminology of obstetric records. Having 20% of the charts abstracted a second time by a different trained research assistant tested inter-rater reliability. Agreement was assessed with the Kappa coefficient, which measures the extent of agreement over and above that which is expected by chance.[5] We considered the Kappa value of 0.78 found for our study sample to be acceptable.[4]

INTERNAL CONSISTENCY

Because the Perinatal Outcome Index measures different aspects of a complex phenomenon, rather than a single concept such as pain or depression, items are not expected to be homogeneous and the inter-item correlation should not be high.[4] We assessed this with Cronbach's alpha, which at 0.55 confirms their heterogeneity.

DISCUSSION

We found the optimality concept an interesting conceptual basis for an index because it defines intrapartum events as "optimal" or "suboptimal", rather than as "normal" or "pathological". The Perinatal Outcome Index provides a summary score of a number of clinical events. We did not include prenatal and post partum information beyond 48 hours in the Canadian version because data are less readily retrievable. We made other modifications to adapt the index to a Canadian setting. We demonstrated that the information required for the Perinatal Outcome Index is readily obtainable from women's health records. The high inter-rater reliability shows that abstraction of data by trained assistants can be accomplished with minimal error. The internal consistency (homogeneity) is low, as desired in a clinimetric index.→

continued from page 13...

The background profile of the 324 women who were included in this study shows that there was a range of differences in social and obstetric variables even though the group was considered to be obstetrically at low risk. Among this study sample, the index scores varied considerably although a large proportion of the sample had scores of 35-36. The clinical importance of differences in scores will become apparent as experience with the use of the index accumulates.

CONCLUSIONS

We believe that this initial assessment of the Perinatal Outcome Index is a step toward establishing its validity and reliability. Our study of its properties has helped to establish its usefulness in a Canadian setting. Additional studies using the index are needed, but it appears to be a useful expression of differences among low risk populations. It is easy to use and data are readily retrievable from both hospital and home birth records. These properties make the index an attractive summary measure for comparative maternity care policy studies, and for quality assurance programs.

It is not designed to evaluate what should or should not have happened during the birth. Interventions may be appropriate but the need for intervention represents a departure from the "optimum". It is a numerical representation of what actually occurred. It combines process and intermediate outcomes into a summary score. We plan to use it as part of a study to compare regulated midwifery care to medical care in Ontario.

ACKNOWLEDGEMENTS

This study was supported by a research grant from the Hamilton Health Sciences Foundation.

Special thanks to the panel of experts: Fionnella S Crombie, MD, CCFP, Associate Clinical Professor, Department of Family Medicine, McMaster University; Vicki Van Wagner, RM, Associate Professor, Ryerson Polytechnic University, Midwifery Education Programme; Barbara Brennan, MD, FRCSC, Associate Professor, Department of Obstetrics and Gynaecology, McMaster University; Lorraine Rodrigues, RN, St. Joseph's Hospital; and Lynn Patterson, consumer.

AUTHORS' BIOGRAPHIES

Patricia McNiven RM, BSc, MSc, PhD is an Assistant Professor in the Department of Family Medicine, Midwifery Education Programme, McMaster University. She is also a practicing midwife in Hamilton, Ontario.

Karyn Kaufman RM, BSN, MS, DrPH is a Professor and Assistant Dean, Faculty of Health Sciences, Midwifery Education Programme, McMaster University. She is also a practicing midwife in Hamilton, Ontario.

Murray Enkin MD, FRCSC, FACOG, LLD, DSc is Professor Emeritus, Department of Obstetrics and Gynecology and Clinical Epidemiology & Biostatistics, McMaster University. He is well known as one of the authors of "A Guide to Effective Care in Pregnancy and Childbirth".

REFERENCES

1. Wiegers TA, Keirse, MJNC, Bergs GAH, van dwe Zee J. An approach to measuring quality of midwifery care. J Clin Epidemiology, 1996;49:(3):319-25.
2. Prechtl HFR. Neurological findings in newborn infants after pre- and perinatal complications. In: Jonxis JHP, Visser HKA, Troelstra JA, editors. Aspects of prematurity and dysmaturity. Leiden: Stenfert Kroese, 1968. p.303-21.
3. Prechtl HFR. The optimality concept. Early Human Development, 1980;(4):210-5.
4. Feinstein AR. Clinimetrics. New Haven: Yale University Press, 1987.
5. Streiner DL, Norman GR. Health measurement scales: a practical guide to their development and use. Oxford: Oxford University Press, 1991.
6. Pedhazur EJ. Multiple regression in behavioural research. 2nd ed. New York: Holt, Rinehart & Winston, 1982.