

Association Not Causation: What Is the Intervention?

Michael C. Klein, CCFP, FAAP, FCFP

Emeritus Professor, Departments of Family Practice and Pediatrics, University of British Columbia, Vancouver, Children's and Women's Health Centre of British Columbia, Vancouver, BC

Ann Fam Med 2007;5:294-297. DOI: 10.1370/afm.729.

Nicholson and colleagues have presented an intriguing and challenging study of a new approach to the management of labor near term,¹ but they fail to convince that their method is actually the reason for their rather astonishing and counterintuitive results. Moreover, there is a danger that others less careful in the application of the study intervention will too enthusiastically embrace the approach—to the detriment of the women and babies whose care the investigators wish to improve.

The investigators applied preventive use of induction of labor for women at or near 38 weeks' gestation based on a scoring tool designed principally to avoid babies that are too large and placentas that are too old to deliver adequate oxygen and nutrition to the fetus (Active Management of Risk in Pregnancy at Term, AMOR-IPAT). The physician is the unit of analysis rather than the woman or the procedure. The investigators achieved enviable results, and though their title speaks of association, it is clear that they believe that their intervention is linked causally to their results.

The problem that the new approach seeks to ameliorate is the unacceptably high cesarean section rates found in their and most other settings. Innovative methods to address this problem are much needed. The literature is replete with quality improvement programs designed to reduce cesarean delivery rates, but it is clear from the general medical literature, as well

as from the obstetric literature pertinent to cesarean section,² that only multiple concurrent approaches are successful, and the magnitude of the resultant change is small.³ Single interventions almost always fail.

NOT A SIMPLE PROBLEM, NOT A SIMPLE SOLUTION

Childbirth is a complex process that begins antepartum, even preconceptionally. Complexity theory would dictate that a simple problem can be solved with a simple solution. If childbirth were simple, a simple intervention such as preventive induction might do the job. Unfortunately, childbirth is not simple, and neither is the proposed intervention.

We can learn something about complex care for a complex problem/phenomenon from midwifery. In midwifery care in the hospital or at home, cesarean section rates and other outcomes are as good as or better than those reported in the current study.⁴ What is the intervention in midwifery care? It is the totality of the care. It includes a coherent philosophy and an articulated, highly noninterventive approach. Induction rates are low, and women whose pregnancy is post-term are usually allowed to progress to labor on their own much later than in conventional physician practice.

Despite their claim that the study physicians' practice style is not different from that of the usual care group, in my view, the study physicians who practice AMOR-IPAT are using a different style. They are a dedicated group with a coherent philosophy who practice an intimate and engaged style of care that is not representative of usual care. They follow the lead of the first author and are convinced that they have a major new approach to an important problem. Although they are not midwives, they practice in many ways like midwives do. This practice style allows them to use an apparently simple intervention such as AMOR-IPAT

Conflict of interest: none reported

CORRESPONDING AUTHOR

Michael C. Klein, CCFP, FAAP, FCFP
Centre for Community and Child Health Research
Child and Family Research Institute, Room L408
4500 Oak St
Vancouver, BC, V6H 3N1 Canada
mklein@interchange.ubc.ca

and get good maternal results without damage to the fetus. In reality, however, their approach is not simple. It is complex and based on the totality of what they do, not only on the 2 elements of AMOR-IPAT.

An historical example of enthusiasts taking up the challenge of their leader before the intervention was fully examined is what happened when Joseph B. DeLee exhorted his followers to implement the "prophylactic forceps operation," with associated episiotomy.⁵ He believed this delivery style could improve outcomes for both mother and fetus at a time when, as today, there was a real problem, when childbirth was indeed quite dangerous. It seemed like a good idea at the time, and it has taken us more than 80 years to recover.

Another example of what happens when there is widespread adoption of an intervention that showed benefit, or at least no apparent problems, also comes from the induction literature. The Canadian Post-Term Trial⁶ and other studies of post-term pregnancy showed no increase in the cesarean section rate and good newborn outcomes when induction is undertaken early in the 41st week of pregnancy.^{7,8} Outside the rarified and controlled conditions of trials, however, early induction for post-term pregnancy leads to a twofold increase in cesarean section rate.⁹ In our Vancouver setting, post-term nulliparous women in spontaneous labor have a cesarean section rate of about 8%, whereas those induced have a cesarean section rate of 44%. Even randomized controlled trials (RCTs) cannot provide information that is externally valid for all settings.

What is the evidence that AMOR-IPAT physicians are different, not just that they used the study intervention?

- The lead author is a member of the study group, which in itself could create bias, and there are other potential bias issues that can result from lack of blinding in the creation and coding of data in such a retrospective study.

- The AMOR-IPAT group is composed of 4 family physicians and 1 obstetrician. The usual care group is composed of 3 family physicians, 1 obstetrician, and a large group of 4 obstetricians and 2 nurse-midwives who share attending deliveries. Thus there are actually 15 clinicians in the study, not 10 as indicated. The AMOR-IPAT physicians attend somewhat more of their deliveries than the usual care group, but it is misleading to analyze the 6 members of the large usual care group as if they were 1 physician. Had the actual individual personal delivery rates of the 6 members of the large usual care group been properly reported, the proportion of births attended by the individual members of the usual care group would drop precipitously, as such groups function on a fixed rotation for deliveries (probably about 1 in 6). Because the large obstetric group

attends 54% of the usual care group births, this loss of intimacy in the comparison group would be large. In fact, I estimate the delivery attendance rates for women that each of the 6 followed antepartum would drop from the reported 96.9% to as low as 16% to 20%.

- The AMOR-IPAT group maintains a substantially lower epidural rate than the usual care group. To do so, care has to be intimate and engaged. Compared with the usual care group, AMOR-IPAT physicians' patients had a longer time from admission to delivery but a similar length of first stage of labor. To achieve this outcome would require the physicians to spend more time with their patients (data not provided). No amount of cluster analysis or other statistical manipulations can explain away this clear practice-style difference. It would have been useful to know the timing of epidurals in the AMOR-IPAT group compared with the usual care group. Although the current Cochrane Collaboration claims epidural analgesia does not increase the cesarean section rate,¹⁰ a sensitivity analysis of early vs late (in active labor) epidurals, shows that early epidural usage can increase the cesarean section rate by more than 2.5 times.¹¹ Based on their intimate style of care, I would predict that the AMOR-IPAT group not only uses epidurals less often but uses them later than the usual care group.

Although the 5-minute APPGAR scores were comparable, the AMOR-IPAT group had statistically fewer neonatal intensive care unit (NICU) admissions. This finding is not meaningful. NICU admissions are likely an artifact subject to local practice differences. Obstetricians often or usually require a pediatrician to be in attendance at birth, and pediatricians tend to bring babies to the NICU, where they can observe, care, and bill for them. This outcome is especially true for babies delivered by cesarean section, and there were many more cesarean section deliveries in the usual care group. Family physicians usually manage the baby themselves, and they have a predisposition for keeping mother and baby together.

RULES OF CAUSALITY

Perhaps one of the best ways of assessing the likelihood that AMOR-IPAT is the actual intervention that makes the difference is to apply Hill's criteria for causality.¹²

- Criterion 1 is met. The investigators have shown a *temporal relationship* between the exposure and the outcome.

- Criterion 2, *size*, is met. They have shown a large difference in outcome between AMOR-IPAT and usual care physicians.

- Criterion 3 is not met. The investigators have not shown a *dose-response relationship*. There are only

2 groups and the cutoffs between the AMOR-IPAT and usual care groups are rather arbitrarily created. If a third or intermediate group had been created, and that group had intermediate results, the causal relation would have been more compelling.

- Criterion 4, *consistency*, is not met. Think of the many studies that were needed before Sir Richard Doll, working with Hill, was able to show that smoking increased the risk (but did not cause) lung cancer.¹³ If the current investigators wish us to believe the association, let alone causal link, between the intervention and the result, they would need to show that their approach gets better results, not only the current study but in many other settings and with many other populations as well. To their credit, their first published study was with an urban population,¹⁴ whereas this study is of a rural population. In the former, however, they used the patient/procedure, not the physician, as the unit of analysis. To show that AMOR-IPAT is the total intervention, the investigators will need to apply the methodology with physicians who are not disciples but who are in true equipoise, that is, they have no biases toward AMOR-IPAT. This study has to be done before even considering an RCT. I suggest studying only obstetricians.

- Criterion 5, *plausibility*, is subject to argument. The investigators would say that their approach is, indeed, plausible based on their research to date. If they are right, then rethinking our conventional wisdom about early induction will be required.

- Criterion 6, *consideration of alternate explanations*, is not met. The investigators insist that they do not practice differently from the usual care group, yet they have not tried to test this belief by any methods except statistical manipulations. A very simple sensitivity analysis or stratification could help here. If they removed the high-volume obstetrician from the AMOR-IPAT-exposed group and the large obstetric group from the nonexposed or usual practice group, power calculations aside, I wonder whether the study results would hold up

- Criterion 7, *the experiment*, or a trial or equivalent has not taken place.

- Criterion 8, *specificity*, is not met. It relates to the usual situation in medicine and biology—that simple solutions are almost always wrong. See the complexity discussion above.

- Criterion 9, *coherence*, is not met. If we believe the authors, then we must reject some of our beliefs about the dangers associated with induction. If the authors are correct, a complete paradigm shift¹⁵ will be required. We are not there, yet. Much more complex experiments will have to take place, and more compelling data will have to be acquired.

OTHER KEY POINTS

Though various statistical tests were used to minimize these differences, demographic analysis shows that women whose physicians were in the AMOR-IPAT group were younger, less often had private medical insurance, much more often had a family practice attending clinician, were more often multiparous, were taller and thinner, and less often had malposition in labor (probably because of less epidural use),¹⁶⁻¹⁸ all of which favor decreases in cesarean rates. Fewer late decelerations were reported on the AMOR-IPAT side, but this finding was perhaps due to less use of continuous electronic fetal monitoring by AMOR-IPAT (data not supplied).

Although there was less blood loss in the AMOR-IPAT patients, this measure is subjective and not supported by either fewer transfusions (data not supplied) or by postpartum hemoglobin levels, which were comparable for both groups. Furthermore, the use of carboprost tromethamine (Hemabate) was similar.

Of interest, in the same 2004 issue as the Nicholson et al first study,¹⁴ the editorialist compared their results with a study published concurrently in the same issue.⁹ That study came to the conventional conclusion that induction raises the cesarean section rate. The editorialist cautioned easy acceptance of AMOR-IPAT, counseling the reader to remember “first do no harm.”¹⁹

In summary, the developers of AMOR-IPAT fail to convince that their intervention is the reason for the outcomes favoring their approach. Many confounders, demographic imbalances, and biases are at play, and the established Hill criteria for causality are far from being met. If this approach were to be taken up by physicians less committed to intimate and personalized care than the investigators and their disciples, our already industrialized maternity care patterns would be augmented by yet another reason to consider birth as an accident waiting to happen, and women would have yet another reason to consider themselves incompetent to give birth without massive intervention.

To read or post commentaries in response to this article, see it online at <http://www.annfamned.org/cgi/content/full/5/4/294>.

Key words: Parturition; cesarean section; preventive care; induction of labor; term birth; epidural analgesia; birth outcomes; maternity care; practice-based research

Submitted June 8, 2007; accepted June 13, 2008.

References

1. Nicholson JM, Yeager D, Macones G. A preventive approach to obstetric care in a rural hospital: association between higher rates of preventive labor induction and lower rates of cesarean delivery. *Ann Fam Med*. 2007;5(4):310-319.

2. Chaillet N, Dube E, Dugas M, et al. Evidence-based strategies for implementing guidelines in obstetrics: a systematic review. *Obstet Gynecol.* 2006;108(5):1234-1245.
3. Flamm BL, Berwick DM, Kabcenell A. Reducing cesarean section rates safely: lessons from a "breakthrough series" collaborative. *Birth.* 1998;25(2):117-124.
4. Janssen PA, Lee SK, Ryan EM, et al. Outcomes of planned home births versus planned hospital births after regulation of midwifery in British Columbia. *CMAJ.* 2002;166(3):315-323.
5. DeLee J. The prophylactic forceps operation. *Am J Obstet Gynecol.* 1920;1:34-44.
6. Hannah ME, Hannah WJ, Hellmann J, Hewson S, Milner R, Willan A. Induction of labor as compared with serial antenatal monitoring in post-term pregnancy. A randomized controlled trial. The Canadian Multicenter Post-term Pregnancy Trial Group. *N Engl J Med.* 1992;326(24):1587-1592.
7. Sanchez-Ramos L, Olivier F, Delke I, Kaunitz AM. Labor induction versus expectant management for postterm pregnancies: a systematic review with meta-analysis. *Obstet Gynecol.* 2003;101(6):1312-1318.
8. Crowley P. Interventions for preventing or improving the outcome of delivery at or beyond term. *Cochrane Database Syst Rev.* 2000;(2): CD000170.
9. Luthy DA, Malmgren JA, Zingheim RW. Cesarean delivery after elective induction in nulliparous women: the physician effect. *Am J Obstet Gynecol.* 2004;191(5):1511-1515.
10. Anim-Somuah M, Smyth R, Howell C. Epidural versus non-epidural or no analgesia in labour. *Cochrane Database Syst Rev.* 2005;(4): CD000331.
11. Klein MC. Does epidural analgesia increase rate of cesarean section? *Can Fam Physician.* 2006;52:419-421, 426-418.
12. Hill AB. The environment and disease: association or causation? *Proc R Soc Med.* 1965;58:295-300.
13. Doll R, Hill AB. Mortality in relation to smoking: ten years' observations of British doctors. *Br Med J.* 1964;1(5396):1460-1467 CONCL.
14. Nicholson JM, Kellar LC, Cronholm PF, Macones GA. Active management of risk in pregnancy at term in an urban population: an association between a higher induction of labor rate and a lower cesarean delivery rate. *Am J Obstet Gynecol.* 2004;191(5):1516-1528.
15. Kuhn T. *The Structure of Scientific Revolutions.* Chicago, Ill: University of Chicago Press; 1970.
16. Hoult IJ, MacLennan AH, Carrie LE. Lumbar epidural analgesia in labour: relation to fetal malposition and instrumental delivery. *Br Med J.* 1977;1(6052):14-16.
17. Yancey MK, Zhang J, Schweitzer DL, Schwarz J, Klebanoff MA. Epidural analgesia and fetal head malposition at vaginal delivery. *Obstet Gynecol.* 2001;97(4):608-612.
18. Klein MC, Grzybowski S, Harris S, et al. Epidural analgesia use as a marker for physician approach to birth: implications for maternal and newborn outcomes. *Birth.* 2001;28(4):243-248.
19. Socol M. Elective induction of labor: should we make a fuss? *Am J Obstet Gynecol.* 2004;191(5):1509-1510.