Association of a quality improvement program with neonatal outcomes in extremely preterm infants: a prospective cohort study

Shoo K. Lee MBBS PhD, Prakesh S. Shah MD MSc, Nalini Singhal MBBS, Khalid Aziz MBBS, Anne Synnes MDCM MHSc, Douglas McMillan MD, Mary M. Seshia MBChB; for the Canadian EPIQ Study Group*

ABSTRACT

Background: We previously demonstrated improvement in bronchopulmonary dysplasia and nosocomial infection among preterm infants at 12 neonatal units using the Evidence-based Practice for Improving Quality (EPIQ). In the current study, we assessed the association of Canada-wide implementation of EPIQ with mortality and morbidity among preterm infants less than 29 weeks gestational age.

Methods: This prospective cohort study included 6026 infants admitted to 25 Canadian units between 2008 and 2012 (baseline year, *n* = 1422; year 1, *n* = 1611; year 2, *n* = 1508; year 3, *n* = 1485). Following a 1-year baseline period and 6 months of training and planning, EPIQ was implemented over 3 years. Our primary outcome was a composite of neonatal mortality and any of bronchopulmonary dysplasia, severe neurologic injury, severe retinopathy of prematurity, necrotizing enterocolitis and nosocomial infection. We compared outcomes for baseline and year 3 using multivariable analyses.

Results: In adjusted analyses comparing baseline with year 3, the composite outcome (70% v. 65%; adjusted odds ratio [OR] 0.63, 95% confidence interval [CI] 0.51 to 0.79), severe retinopathy (17% v. 13%; OR 0.60, 95% CI 0.45 to 0.79), necrotizing enterocolitis (10% v. 8%; OR 0.73, 95% CI 0.52 to 0.98) and nosocomial infections (32% v. 24%; OR 0.63, 95% CI 0.48 to 0.82) were significantly reduced. The composite outcome was lower among infants born at 26 to 28 weeks gestation (62% v. 52%; OR 0.62, 95% CI 0.49 to 0.78) but not among infants born at less than 26 weeks gestational age (90% v. 88%; OR 0.73, 95% CI 0.44 to 1.20).

Interpretation: EPIQ methodology was generalizable within Canada and was associated with significantly lower likelihood of the composite outcome, severe retinopathy, necrotizing enterocolitis and nosocomial infections. Infants born at 26 to 28 weeks gestational age benefited the most.

Rates of neonatal mortality and morbidity have not improved substantially over the past decade, despite advances in neonatal and perinatal care.[1–3] To improve infant outcomes, quality improvement initiatives have been implemented in neonatal units, but their success has varied.[4–6] A common approach is to form collaborative neonatal networks and to use quality improvement methods to address variations in practice and outcomes.[7–9] Commonly used quality improvement methods include identifying the best available evidence, adapting practice change to the organization, collaborative learning, and audit and feedback of information implemented in iterative "plan, do, study, act" cycles.[10]

We previously developed and tested the Evidence-based Practice for Improving Quality (EPIQ) in 12 neonatal units in a cluster randomized trial.[11] EPIQ is a multifaceted quality improvement approach that combines the best available evidence with institution-specific data to identify institution-specific needs. Sites targeting nosocomial infection achieved a 32% reduction in infections, and sites targeting bronchopulmonary dysplasia achieved a 15% reduction in that condition. Five units that did not participate formed a comparison group and showed no improvement for either outcome.[11]

In the current study, our objective was to assess the generalizability of EPIQ to other neonatal units in Canada and to determine any association with improvements in mortality and morbidity among preterm neonates born at less than 29 weeks gestational age.

Methods

This prospective cohort study involved 25 of the 30 tertiary-level neonatal units within Canada that participate in the Canadian Neonatal Network and included 95% of all eligible neonatal admissions to those units. Of the remaining units, 3 were small, with insufficient resources to implement EPIQ, and 2 were not fully accredited as tertiary-level units at the time of study commencement. We included all infants less than 29 weeks gestational age who were admitted to the participating units, except those for whom palliative care was planned at birth. The 25 participating sites had a median of 26 (interquartile range [IQR] 14–36) beds and admitted a median of 69 (IQR 31–128) neonates per year during the period 2008 to 2012. The 12 sites from the initial EPIQ trial were included in this project. Infants from the 13 additional sites had higher mean birth weight (971 v. 924 g), and a lower proportion had an Apgar score below 7 (36% v. 42%). We calculated gestational age using a hierarchical algorithm ordered as date of in vitro fertilization, early ultrasound, last menstrual period, obstetric estimate and pediatric estimate.

Conducted from April 2008 to September 2012, the study consisted of a 1-year period for collecting baseline data, a 6-month quality improvement training and planning period, and a 3-year intervention period. Participating units chose the outcomes to be targeted and the practice changes to be implemented. All units were aware of the others' activities via a web portal.[12] The study was approved by the institutions' ethics review boards. Informed consent from patients was waived, as the activities were considered part of routine quality improvement. The data monitoring committee recommended continuation of the study on an annual basis.

Outcomes

The primary outcome was a composite of death or any of 5 major morbidities: bronchopulmonary dysplasia (need for supplemental oxygen at 36 weeks corrected gestational age or time of transfer to a level-2 facility[13]), severe neurologic injury (grade ≥ 3 intraventricular hemorrhage[14] or intraparenchymal echogenicity or periventricular leukomalacia), severe retinopathy of prematurity (stage ≥ 2[15] or need for treatment in either eye), necrotizing enterocolitis (stage ≥ 2[16]) and nosocomial infection (positive results on culture of blood or cerebrospinal fluid samples obtained after 2 days of age from a symptomatic neonate[17]). We also examined these outcomes independently as secondary outcomes. The oxygen reduction test was not used for diagnosing bronchopulmonary dysplasia. The numbers of infants who underwent neurologic imaging or an eye examination were less than the total number of infants included in the study. All data were abstracted from patient charts according to standardized definitions[18] and were transmitted electronically to the Coordinating Centre.

Baseline data collection, quality improvement training and planning

During the baseline period, we conducted interviews, focus groups, group sessions and staff surveys at each unit to evaluate organizational culture and behaviour. We reported results back to the units, along with our recommendations. Site investigators established EPIQ teams, which typically included a neonatologist, nurse manager, nurse educator, quality-improvement personnel and outcome-specific personnel. Site investigators and 2 or 3 team members from each unit attended training and a planning workshop.

Training covered use of Pareto charts, histograms, control charts, trend analysis, rapid-cycle improvement models[19] and change management techniques. We emphasized the value of identifying nonambiguous clinical aims, team-building and leadership, defragmenting the process to identify barriers, developing flow charts of existing and "ideal" processes, and generating feedback. Participants then trained their hospital teams. The planning workshop consisted of group sessions to identify best practices. Each group created lists of interventions that were known or anticipated to improve a specific outcome and then performed systematic reviews using a rapid review format. Unit teams evaluated their own baseline data, selected outcomes to target and developed process measures.

Information on proposed interventions, plans, dissemination materials and process measures were shared among the teams. Toward the end of the baseline training period, outcome groups reviewed the accumulated evidence, data analysis, clinical process analyses and qualitative study reports. We classified the interventions into 3 categories according to the evidence (definitive, nondefinitive and unsubstantiated) and established consensus recommendations. Evidence and recommendations were continually reviewed throughout the study. Training and planning workshops were repeated annually.

Interventions

Participating units implemented practice changes every 3 to 6 months using data from process indicators, quarterly benchmarking reports and consensus recommendations, according to local organizational and cultural issues. We monitored

compliance at random intervals. At the annual workshops, each site reported its successes, failures and barriers to implementing change. Top-performing sites for each outcome shared their approaches. Discussion and opportunities for mutual learning were facilitated. We provided quarterly feedback to the sites regarding their outcomes and process performance in comparison with the network. Midway through the intervention period, site representatives were encouraged to visit other sites to observe practices and share change strategies.

Statistical analyses

We made an a priori decision to compare the change in rate of the composite outcome between baseline and year 3 of the intervention period (for the power calculation, see Appendix 2, available at www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.140399/-/DC1). We summarized characteristics of the study population using descriptive statistics. We compared infant characteristics among the baseline and intervention years using the χ^2 test for categorical measures and the F test or Kruskal–Wallis test, as appropriate, for continuous measures. We examined homogeneity of the intervention effects across hospitals using a forest plot of hospital-specific odds ratios (ORs) with a summary OR estimated using a random-effects model. When significant change in an adverse outcome was observed, we estimated number needed to treat and 95% confidence intervals (CIs) from raw outcome rates without adjustment. To account for site effects and the heterogeneity of intervention effects across sites, we applied hierarchical logistic regression models[20] with random intercept and random site-intervention interaction adjusting for potential confounders (gestational age, small for gestational age, Apgar score < 7 at 5 minutes, Score for Neonatal Acute Physiology version II[21] > 20) to examine the change in rate of outcome from baseline to each intervention year. We used the compound symmetric covariance structure for the regressions. We performed stratified analyses for groups with gestational age less than 26 weeks and gestational age 26 to 28 weeks. For the entire cohort and the subgroups with gestational age less than 26 weeks and 26 to 28 weeks, gestational age was used as a continuous variable.

We conducted exploratory analyses to assess the effects of specific interventions and their association with outcome from sites that reported implementing each practice change. For each intervention, we compared pre- and post-implementation data using unadjusted analyses. We used SAS version 9.3 and R version 2.15.1 software for data management and statistical analyses. We applied a significance level of 0.05 for 2-sided tests, without multiple-comparison adjustment, as these analyses were complementary to the primary planned analyses.

Results

The 25 sites implemented a total of 130 practice changes targeting a specific outcome and 30 interventions that improved the care process. The number of practice changes implemented per site ranged from 2 to 8 and were grouped by outcome (Table 1 and Appendix 2); however, some changes

Table 1: Practice changes implemented by participating sites during the intervention years	
Practice change groups, categorized by outcome	No. of sites
Bronchopulmonary dysplasia	
Use of T-piece resuscitator	7
Enhancing or ensuring early use of continuous positive airway pressure	5
Controlled oxygen use or review of oxygen saturation	5
Early use of surfactant and/or early extubation	5
Ventilator weaning algorithm	3
Prophylactic surfactant	1
Severe neurologic injury	
Delayed cord clamping	7
Use of antenatal magnesium sulphate	4
Minimal use of volume expanders	1
Minimal use of inotropes	1
Retinopathy of prematurity	
Controlled oxygen use	11
Improved screening protocol	1
Necrotizing enterocolitis	
Feeding guidelines	11
Early feeding	5
Use of donor milk	4
Colostrum or enhanced expressed breast milk	3
Early total parenteral nutrition	2
Hold enteral feeds during erythrocyte transfusion	2
Nosocomial infection	
Central line management, bundle	13
Hand hygiene	10
Central line management, single intervention only	9
Education, audit, team-building	6
Reduce skin breaks	5
Skin-to-skin care	4
Checklist for central lines	3

could affect multiple outcomes. The most commonly adopted changes were use of a T-piece resuscitator, early and adequate use of continuous positive airway pressure, controlled oxygen use, and early use of surfactant and/or early extubation for bronchopulmonary dysplasia; delayed cord clamping for neurologic injury; controlled oxygen use for retinopathy of prematurity; standardization of feeding guidelines and donor milk for necrotizing enterocolitis; and hand hygiene and central line management practices for nosocomial infection.

The proportions of infants admitted to units at less than 26 weeks gestation, with Apgar score

less than 7 at 5 minutes, and with Score for Neonatal Acute Physiology version II above 20 were greater in year 3 than in the baseline year (Table 2). Process measures confirmed implementation of various practice changes (e.g., decreased use of mechanical ventilation, older age at first administration of surfactant, increased use of magnesium sulphate).

Compared with baseline, there was a reduction in the composite outcome in year 3, as well as in severe retinopathy and nosocomial infections (Table 3). Mild heterogeneity ($I^2 = 44\%$) was observed among site-specific ORs for the

Table 2: Infant and maternal characteristics and process measures

Characteristic	Study year; no. (%) of infants*				p value	
	Baseline $n = 1422$	Year 1 $n = 1611$	Year 2 $n = 1508$	Year 3 $n = 1485$	Overall†	Baseline v. yr 3‡
Sex, male	778 (55)	897 (56)	799 (53)	841 (57)	0.2	0.3
Gestational age (GA)					< 0.001	< 0.001
22^0 to 25^6 wk§	403 (28)	519 (32)	477 (32)	534 (36)		
26^0 to 28^6 wk§	1019 (72)	1092 (68)	1031 (68)	951 (64)		
Mean ± SD, wk	26.3 ± 1.5	26.2 ± 1.5	26.2 ± 1.5	26.1 ± 1.5	< 0.001	< 0.001
Birth weight, g, mean ± SD	947 ± 274	920 ± 238	919 ± 246	919 ± 256	0.004	0.004
Small for GA	130 (9)	133 (8)	121 (8)	19 (8)	0.6	0.3
Outborn¶	304 (21)	304 (19)	284 (19)	283 (19)	0.2	0.1
Apgar score < 7 at 5 min	543 (38)	618 (38)	628 (42)	675 (45)	< 0.001	< 0.001
SNAP-II score > 20	393 (28)	475 (29)	457 (30)	476 (32)	0.05	0.006
Singleton	1012 (71)	1153 (72)	1090 (72)	1089 (73)	0.6	0.2
Cesarean birth	801 (56)	910 (56)	842 (56)	833 (56)	0.9	0.6
Maternal hypertension**	203/1391 (15)	245/1560 (16)	238/1483 (16)	211/1436 (15)	0.6	0.9
Maternal diabetes mellitus**	96/1389 (7)	106/1531 (7)	116/1454 (8)	108/1415 (8)	0.6	0.5
Antenatal steroids**	1151/1374 (84)	1358/1562 (87)	1264/1465 (86)	1258/1458 (86)	0.08	0.06
Intrapartum magnesium sulphate**	NA	194/1133 (17)	320/1437 (22)	582/1419 (41)	< 0.001	NA
Surfactant use	1115 (78)	1271 (79)	1182 (78)	1126 (76)	0.2	0.1
Age at first surfactant, median (IQR), min	31 (15–93)	39 (17–161)	44 (19–201)	46 (20–212)	< 0.001	0.02
Mechanical ventilation, any time	1260 (89)	1391 (86)	1268 (84)	1228 (83)	< 0.001	< 0.001

Note: EPIQ = Evidence-based Practice for Improving Quality, IQR = interquartile range, NA = not available, SD = standard deviation, SNAP-II = Score for Neonatal Acute Physiology, version II (a measure of severity of illness on admission to neonatal intensive care unit; scores range from 0 to 115, with higher scores indicating increasing sickness).[21]
*Except where indicated otherwise.
†Calculated using χ^2 test for categorical variables and F test or Wilcoxon rank-sum test, as appropriate, for continuous variables.
‡Based on comparison between baseline and year 3.
§Superscript 0 and 6 denote number of days in addition to weeks of gestation.
¶Born at a hospital other than the institution that housed the admitting neonatal intensive care unit.
**Data missing for some infants; denominators used to calculate percentages (shown in these rows) reflect available data.

composite outcome (Figure 1). The summary OR, estimated using a random-effects model, was 0.74 (95% CI 0.59 to 0.94). After multivariable analyses, we observed statistically significant reductions in year 3 compared with baseline for the composite outcome (OR 0.63, 95% CI 0.51 to 0.79; number needed to treat to prevent 1 incident [NNT] 18, 95% CI 12 to 45), severe retinopathy (OR 0.60, 95% CI 0.45 to 0.79; NNT 26, 95% CI 14 to 133), necrotizing enterocolitis (OR 0.73, 95% CI 0.52 to 0.98; NNT 50, 95% CI 25 to 714) and nosocomial infection (OR 0.63, 95% CI 0.48 to 0.82; NNT 14, 95% CI 10 to 25) (Table 4). No statistically significant differences in mortality (OR 0.86, 95% CI 0.65 to 1.13), bronchopulmonary dysplasia (OR 0.81,

95% CI 0.64 to 1.01) or severe neurologic injury (OR 0.91, 95% CI 0.69 to 1.19) were observed. The majority of improvements were observed in infants born at 26 to 28 weeks gestational age. Improvement in outcomes accelerated as the study progressed, with maximum benefit in year 3 (Table 4 and Appendix 2).

Table 5 shows the association of specific practice changes with outcomes using site-specific analyses. Three practice changes were identified as significant contributors to outcomes change: early use of surfactant and/or early extubation was associated with a decrease of 17.7 (95% CI 10.2 to 25.0) percentage points in bronchopulmonary dysplasia, use of colostrum or enhanced breast milk with a decrease of 5.3

Table 3: Neonatal outcomes during study period

Outcome*	Study year; no. affected/no. in group (%)			
	Baseline $n = 1422$	Year 1 $n = 1611$	Year 2 $n = 1508$	Year 3 $n = 1485$
Composite	997/1422 (70)	1106/1611 (69)	1065/1508 (71)	958/1485 (65)
GA 22^0 to 25^6 wk	364/403 (90)	462/519 (89)	434/477 (91)	468/534 (88)
GA 26^0 to 28^6 wk	633/1019 (62)	644/1092 (59)	631/1031 (61)	490/951 (52)
Death	219/1422 (15)	257/1611 (16)	242/1508 (16)	249/1485 (17)
GA 22^0 to 25^6 wk	128/403 (32)	165/519 (32)	153/477 (32)	166/534 (31)
GA 26^0 to 28^6 wk	91/1019 (9)	92/1092 (8)	89/1031 (9)	83/951 (9)
Bronchopulmonary dysplasia	483/1200 (40)	578/1345 (43)	564/1268 (44)	486/1245 (39)
GA 22^0 to 25^6 wk	164/281 (58)	230/352 (65)	215/323 (67)	221/373 (59)
GA 26^0 to 28^6 wk	319/919 (35)	348/993 (35)	349/945 (37)	265/872 (30)
Severe neurologic injury†	272/1274 (21)	301/1486 (20)	289/1437 (20)	305/1412 (22)
GA 22^0 to 25^6 wk	137/355 (39)	152/462 (33)	139/448 (31)	177/498 (36)
GA 26^0 to 28^6 wk	135/919 (15)	149/1024 (15)	150/989 (15)	128/914 (14)
Severe retinopathy of prematurity†	165/983 (17)	160/1120 (14)	169/1075 (16)	139/1073 (13)
GA 22^0 to 25^6 wk	109/271 (40)	109/331 (33)	106/304 (35)	107/361 (30)
GA 26^0 to 28^6 wk	56/712 (8)	51/789 (6)	63/771 (8)	32/712 (4)
Necrotizing enterocolitis	133/1390 (10)	146/1604 (9)	130/1508 (9)	113/1485 (8)
GA 22^0 to 25^6 wk	53/388 (14)	59/515 (11)	58/477 (12)	56/534 (10)
GA 26^0 to 28^6 wk	80/1002 (8)	87/1089 (8)	72/1031 (7)	57/951 (6)
Nosocomial infection	451/1422 (32)	452/1611 (28)	407/1508 (27)	363/1485 (24)
GA 22^0 to 25^6 wk	166/403 (41)	205/519 (39)	180/477 (38)	199/534 (37)
GA 26^0 to 28^6 wk	285/1019 (28)	247/1092 (23)	227/1031 (22)	164/951 (17)

Note: GA = gestational age.
*For GA subgroups, superscript 0 and 6 denote number of days in addition to weeks of gestation.
†Among those who underwent examination.

(95% CI 1.9 to 8.9) percentage points in necrotizing enterocolitis and central line management with decreases of 9.3 (single intervention, 95% CI 5.5 to 12.9) and 5.7 (bundled intervention, 95% CI 2.8 to 8.6) percentage points in nosocomial infections. Delayed cord clamping was reported as a practice change by 7 sites; however, of 521 infants in the postimplementation phase, only 166 actually received delayed cord clamping, and of these 22% had severe neurologic injury (same rate as entire cohort).

Interpretation

In this prospective cohort study, embracing the entire EPIQ process was likely key to successful change management and more important than any single intervention. Certain specific interventions were identified as contributing significantly to improvement in outcomes, but the study was underpowered to infer these associations as conclusive. Our finding that targeting all important neonatal outcomes led to success in some outcomes and little change in others was unsurprising, given that units were free to identify their own targets. The lack of a reduction in

mortality may have been due to residual confounding related to higher prematurity and higher illness severity during the later intervention years. In year 3, we observed a reduction in the rate of bronchopulmonary dysplasia; however, continued monitoring is required to determine the sustainability of this change. Delaying the age at which surfactant was administered without a significant reduction in the number of infants requiring surfactant suggests that proper evaluation of change in strategies is warranted.

Although our initial EPIQ trial[11] reported significant reductions in infections and bronchopulmonary dysplasia, we were not able to replicate those results when the trial interventions were offered to all other Canadian units as best practice guidelines.[22] In the later study, continued improvement was observed in the 11 units from the initial EPIQ trial but not the 14 units that were new to EPIQ. This result is consistent with previous observations that practice guidelines do not necessarily result in improvement[23] and that multifaceted programs aimed at changing behaviours are more effective for quality improvement.[24] Thus, a collaborative decision was made to involve all units in this third study of EPIQ.

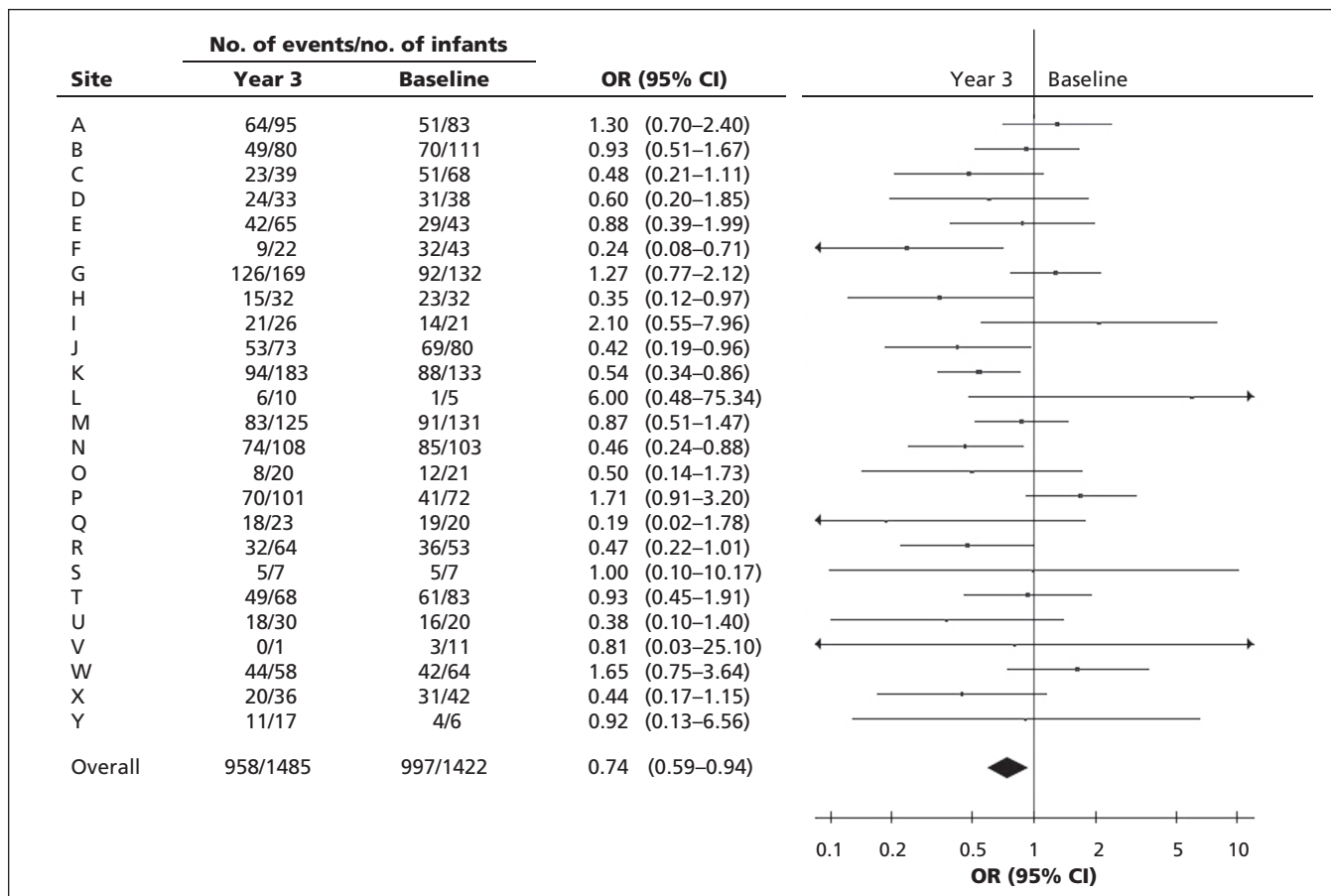| Site	No. of events/no. of infants Year 3	Baseline	OR (95% CI)	Year 3	Baseline
A	64/95	51/83	1.30 (0.70–2.40)		
B	49/80	70/111	0.93 (0.51–1.67)		
C	23/39	51/68	0.48 (0.21–1.11)		
D	24/33	31/38	0.60 (0.20–1.85)		
E	42/65	29/43	0.88 (0.39–1.99)		
F	9/22	32/43	0.24 (0.08–0.71)		
G	126/169	92/132	1.27 (0.77–2.12)		
H	15/32	23/32	0.35 (0.12–0.97)		
I	21/26	14/21	2.10 (0.55–7.96)		
J	53/73	69/80	0.42 (0.19–0.96)		
K	94/183	88/133	0.54 (0.34–0.86)		
L	6/10	1/5	6.00 (0.48–75.34)		
M	83/125	91/131	0.87 (0.51–1.47)		
N	74/108	85/103	0.46 (0.24–0.88)		
O	8/20	12/21	0.50 (0.14–1.73)		
P	70/101	41/72	1.71 (0.91–3.20)		
Q	18/23	19/20	0.19 (0.02–1.78)		
R	32/64	36/53	0.47 (0.22–1.01)		
S	5/7	5/7	1.00 (0.10–10.17)		
T	49/68	61/83	0.93 (0.45–1.91)		
U	18/30	16/20	0.38 (0.10–1.40)		
V	0/1	3/11	0.81 (0.03–25.10)		
W	44/58	42/64	1.65 (0.75–3.64)		
X	20/36	31/42	0.44 (0.17–1.15)		
Y	11/17	4/6	0.92 (0.13–6.56)		
Overall	958/1485	997/1422	0.74 (0.59–0.94)		

Figure 1: Forest plot of odds ratios (ORs) for the composite outcome between baseline and intervention year 3 by site (identifed by letters). The summary OR was calculated using a random-effects model. Statistical heterogeneity (I^2) was 44%. CI = confidence interval.

Our approach allowed us to test the generalizability of EPIQ in a more pragmatic setting, as each unit had different needs, some interventions affected multiple outcomes, and many units routinely targeted multiple outcomes simultaneously. The downside of our approach was that many units targeted outcomes such as nosocomial infections and retinopathy and made little effort to address other outcomes.

Our findings are consistent with reports from other neonatal quality improvement initiatives, such as the Vermont Oxford Network.[25] Although their approach differed from ours in some respects, both were based on evidence review, cultivating quality improvement at a systems level and using team approaches to implement practice changes. The Vermont Oxford Network has reported decreased length of stay and improvements in infant growth,[26,27] severe retinopathy, neurologic injury, oxygen use at discharge[28] and infections.[29] Similarly, the California Perinatal Care Collaborative showed decreased rates of necrotizing enterocolitis,[30] increased administration of antenatal steroids[31] and decreased infections following state-based quality improvement efforts.[6] The difference between these approaches and ours lies in our use of central coordination and continued support; our ability to address multiple interventions for multiple outcomes; unit-led practice changes in our study, rather than central enforcement, which created a culture of autonomy, accountability and responsibility; and annual in-person meetings and training that fostered extensive collaboration.

Limitations

Our study involved many different interventions, and was not powered to identify the individual effects of a specific intervention. We conducted exploratory analyses to address this issue but acknowledge that the current report involves multiple comparisons without statis-

Table 4: Adjusted comparison of neonatal outcomes between baseline and 3 intervention years

Outcome†	Comparison; adjusted OR (95% CI)*		
	Baseline to year 1	Baseline to year 2	Baseline to year 3
Composite	0.87 (0.70 to 1.08)	0.95 (0.76 to 1.18)	0.63 (0.51 to 0.79)
GA 22^0 to 25^6 wk	0.84 (0.51 to 1.39)	0.99 (0.59 to 1.68)	0.73 (0.44 to 1.20)
GA 26^0 to 28^6 wk	0.89 (0.70 to 1.12)	0.94 (0.74 to 1.19)	0.62 (0.49 to 0.78)
Death	0.88 (0.67 to 1.15)	0.83 (0.63 to 1.09)	0.86 (0.65 to 1.13)
GA 22^0 to 25^6 wk	0.86 (0.59 to 1.27)	0.89 (0.60 to 1.30)	0.86 (0.59 to 1.25)
GA 26^0 to 28^6 wk	0.93 (0.59 to 1.44)	0.74 (0.47 to 1.16)	0.87 (0.56 to 1.38)
Bronchopulmonary dysplasia	1.04 (0.83 to 1.44)	1.10 (0.88 to 1.38)	0.81 (0.64 to 1.01)
GA 22^0 to 25^6 wk	1.47 (1.01 to 2.12)	1.41 (0.97 to 2.06)	1.12 (0.78 to 1.61)
GA 26^0 to 28^6 wk	0.96 (0.75 to 1.24)	1.04 (0.80 to 1.34)	0.76 (0.59 to 0.99)
Severe neurologic injury	0.96 (0.73 to 1.25)	0.88 (0.67 to 1.15)	0.91 (0.69 to 1.19)
GA 22^0 to 25^6 wk	0.79 (0.57 to 1.11)	0.71 (0.51 to 0.99)	0.82 (0.59 to 1.14)
GA 26^0 to 28^6 wk	1.08 (0.77 to 1.52)	1.02 (0.72 to 1.44)	0.95 (0.67 to 1.35)
Severe retinopathy of prematurity	0.78 (0.60 to 1.03)	0.88 (0.67 to 1.14)	0.60 (0.45 to 0.79)
GA 22^0 to 25^6 wk	0.72 (0.48 to 1.07)	0.75 (0.50 to 1.12)	0.59 (0.40 to 0.88)
GA 26^0 to 28^6 wk	0.87 (0.58 to 1.33)	1.08 (0.72 to 1.61)	0.57 (0.36 to 0.92)
Necrotizing enterocolitis	0.98 (0.70 to 1.37)	0.83 (0.59 to 1.17)	0.73 (0.52 to 0.98)
GA 22^0 to 25^6 wk	0.86 (0.50 to 1.45)	0.97 (0.57 to 1.64)	0.77 (0.45 to 1.31)
GA 26^0 to 28^6 wk	0.98 (0.70 to 1.37)	0.83 (0.58 to 1.17)	0.72 (0.50 to 1.04)
Nosocomial infection	0.84 (0.64 to 1.08)	0.79 (0.61 to 1.03)	0.63 (0.48 to 0.82)
GA 22^0 to 25^6 wk	0.95 (0.69 to 1.30)	0.89 (0.64 to 1.22)	0.85 (0.62 to 1.16)
GA 26^0 to 28^6 wk	0.78 (0.58 to 1.07)	0.75 (0.55 to 1.02)	0.51 (0.37 to 0.71)

Note: CI = confidence interval, GA = gestational age, OR = odds ratio.
*Adjusted ORs (and 95% CIs) based on hierarchical logistic regression models with random intercept and random site-intervention interaction, adjusted for sex, GA, small-for-GA status, Apgar score at 5 min and Score for Neonatal Acute Physiology, version II (SNAP-II score) > 20.
†For GA subgroups, the superscript 0 and 6 denotes number of days in addition to weeks of gestation.

Table 5 (part 1 of 2): Association of practice changes with neonatal outcomes, according to reported practice changes

Practice changes grouped by outcome	Practice change implemented		Cumulative pre-implementation		Cumulative post-implementation		Outcome rate pre-implementation*		Outcome rate post-implementation*		Absolute change in risk (95% CI)
	No. of sites	No. of infants	Time, mo	No. of infants	Time, mo	No. of infants	%	n/N	%	n/N	
Bronchopulmonary dysplasia											
Use of T-piece resuscitator	7	2386	191	1279	149	1107	46.4 (43.3 to 49.3)	501/1081	46.1 (42.8 to 49.4)	414/898	−0.2 (−4.6 to 4.2)
Enhancing or ensuring early CPAP use	5	1405	167	1129	76	276	44.4 (41.2 to 47.6)	420/947	43.0 (36.6 to 49.4)	98/228	−1.4 (−8.6 to 5.8)
Controlled oxygen use/ review oxygen saturation	5	1437	137	946	106	491	35.9 (32.6 to 39.2)	289/806	31.7 (27.2 to 36.2)	131/413	−4.2 (−9.8 to 1.4)
Early surfactant and/or early extubation	5	820	101	406	142	414	55.1 (49.8 to 60.6)	182/330	37.6 (32.5 to 42.7)	130/346	−17.7 (−25.0 to −10.2)
Ventilator weaning algorithm	3	669	44	181	102	488	52.2 (44.4 to 60.2)	81/155	55.4 (50.6 to 60.0)	238/430	3.1 (−6.2 to 12.2)
Prophylactic surfactant	1	23	12	11	37	12	25.0 (0.0 to 55.0)	2/8	36.4 (8.0 to 64.8)	4/11	11.4 (−29.9 to 52.7)
Severe neurologic injury											
Delayed cord clamping	7	1260	190	739	150	521	19.1 (16.1 to 22.1)	130/682	25.6 (21.8 to 29.4)	127/496	6.5 (1.7 to 11.3)
Use of antenatal $MgSO_4$	4	902	129	545	65	357	20.4 (16.8 to 24.0)	100/491	22.4 (17.9 to 26.9)	73/326	2.0 (−3.8 to 7.8)
Minimal use of volume expanders	1	130	16	46	33	84	25.0 (11.6 to 38.4)	10/40	10.7 (3.7 to 17.7)	8/75	−14.3 (−29.4 to −0.8)
Minimal use of inotropes	1	130	12	32	37	98	35.7 (18.0 to 53.4)	10/28	9.2 (3.1 to 15.3)	8/87	−26.5 (−45.3 to −7.7)
Severe retinopathy of prematurity											
Controlled oxygen use	11	2780	309	1783	225	997	15.1 (13.1 to 17.1)	183/1213	14.2 (11.6 to 16.8)	97/681	−0.9 (−4.2 to 2.4)
Improved screening protocol	1	46	40	32	9	14	23.8 (5.6 to 42.0)	5/21	0	0/7	−23.8 (−42.0 to −5.6)
Necrotizing enterocolitis											
Feeding guidelines	11	2692	292	1372	242	1320	10.0 (8.4 to 11.6)	135/1351	7.9 (6.4 to 9.4)	104/1320	−2.1 (−4.3 to 0.1)
Early feeding	5	1353	110	739	133	614	6.1 (4.4 to 7.8)	45/734	8.8 (6.6 to 11.0)	54/613	2.7 (−0.001 to 5.5)
Use of donor milk	4	971	85	645	61	326	11.0 (8.6 to 13.4)	70/634	8.9 (5.8 to 12.0)	29/325	−2.1 (−6.0 to 1.8)
Colostrum/ enhanced expressed breast milk	3	1140	56	355	90	785	10.2 (7.0 to 13.4)	35/344	4.8 (3.3 to 6.3)	38/785	−5.3 (−8.9 to −1.9)
Early total parenteral nutrition	2	115	52	66	45	49	1.5 (0.0 to 4.5)	1/65	4.1 (0.0 to 9.7)	2/49	2.5 (−3.7 to 8.9)
Hold enteral feeds during erythrocyte transfusions	2	297	70	218	27	79	11.1 (6.9 to 15.3)	24/217	5.1 (0.2 to 10.0)	4/79	−6.0 (−12.4 to 0.4)

Table 5 (part 2 of 2): Association of practice changes with neonatal outcomes, according to reported practice changes

Practice changes grouped by outcome	Practice change implemented		Cumulative pre-implementation		Cumulative post-implementation		Outcome rate pre-implementation*		Outcome rate post-implementation*		Absolute change in risk (95% CI)
	No. of sites	No. of infants	Time, mo	No. of infants	Time, mo	No. of infants	%	n/N	%	n/N	
Nosocomial infection											
Central line management, bundle	13	3859	327	2110	304	1749	31.8 (29.8 to 33.8)	670/ 2110	26.1 (24.0 to 28.2)	457/ 1749	−5.7 (−8.6 to −2.8)
Hand hygiene	10	2350	213	963	272	1387	32.1 (29.2 to 35.0)	309/ 963	30.4 (28.0 to 32.8)	421/ 1381	−1.7 (−5.5 to 2.1)
Central line management, single intervention	9	2294	284	1493	153	801	31.4 (28.9 to 33.7)	468/ 1493	22.1 (19.2 to 25.0)	177/ 801	−9.3 (−12.9 to −5.5)
Education, audit, team-building	6	1505	109	448	183	1057	31.9 (27.6 to 36.2)	143/ 448	30.6 (27.8 to 33.4)	323/ 1057	−1.3 (−6.4 to 3.8)
Reduce skin breaks	5	802	126	372	117	430	31.5 (26.8 to 36.2)	117/ 372	35.8 (31.3 to 40.3)	154/ 430	4.4 (−2.2 to 10.8)
Skin-to-skin care	4	1231	87	505	108	726	32.3 (28.2 to 36.4)	163/ 505	28.1 (24.8 to 31.4)	204/ 726	−4.2 (−9.4 to 1.0)
Checklist for central lines	3	436	81	237	65	199	29.1 (23.3 to 34.9)	69/ 237	27.6 (21.4 to 33.8)	55/ 199	−1.5 (−10.0 to 7.0)

Note: CI = confidence interval, CPAP = continuous positive airway pressure, $MgSO_4$ = magnesium sulphate.
*Outcome rates (and 95% CIs) were calculated according to the date of implementation of the specific change at sites that reported implementing that change, where "pre-implementation" is the period before the date of implementation and "post-implementation" is the period after the date of implementation (within the study period), based on n/N, where n = number of infants with the outcome and N = number of infants with data available for that outcome.

tical adjustment and that the observed improvements may be attributable to overall culture change rather than specific interventions. It is also possible that some interventions took longer than planned to implement, as for delayed cord clamping. In addition, similar to the situation for other observational studies, we were unable to ascertain the effects of secular change, other uncontrolled interventions and changes in management practices on the observed differences in outcomes. Variations in outcomes by site were evident in each quarterly report, and these variations should be explored, along with interactions among multiple interventions. The number of smaller, younger and sicker infants increased as the study progressed, which emphasizes the importance of stratified and adjusted analyses and suggests possible residual confounding related to prematurity and illness severity. However, we observed incremental improvement in outcomes despite increased severity of illness on admission, which suggests that quality improvement initiatives take time to produce results. Variation in the number of interventions by individual units may also have limited the generation of specific intervention–time–outcome relationships. However, this pragmatic approach improves the generalizability of findings.

Conclusion

In this study, we have shown that EPIQ is generalizable within Canada and is associated with a significantly lower likelihood of a composite adverse outcome, severe retinopathy, necrotizing enterocolitis and nosocomial infections among infants born at less than 29 weeks gestational age, with those born at 26 to 28 weeks gestational age benefiting the most. Further research is needed to address outcomes in those born at less than 26 weeks gestational age. Further work to develop guided practice changes, perform accountability audits and generate a higher level of collaboration through mutual site visits is underway, with the launch of a new project in February 2014 that includes all 30 Canadian units.

References

1. Fanaroff AA, Stoll BJ, Wright LL, et al. Trends in neonatal morbidity and mortality for very low birthweight infants. *Am J Obstet Gynecol* 2007;196:147.e1-8.
2. Horbar JD, Carpenter JH, Badger GJ, et al. Mortality and neonatal morbidity among infants 501 to 1500 grams from 2000 to 2009. *Pediatrics* 2012;129:1019-26.
3. Shah PS, Sankaran K, Aziz K, et al. Outcomes of preterm

infants < 29 weeks gestation over 10-year period in Canada: A cause for concern? *J Perinatol* 2012;32:132-8.

4. Horbar JD, Carpenter JH, Buzas J, et al. Collaborative quality improvement to promote evidence based surfactant for preterm infants: a cluster randomised trial. *BMJ* 2004;329:1004.

5. Walsh M, Laptook A, Kazzi SN, et al. A cluster-randomized trial of benchmarking and multimodal quality improvement to improve rates of survival free of bronchopulmonary dysplasia for infants with birth weights of less than 1250 grams. *Pediatrics* 2007;119:876-90.

6. Wirtschafter DD, Powers RJ, Pettit JS, et al. Nosocomial infection reduction in VLBW infants with a statewide quality-improvement model. *Pediatrics* 2011;127:419-26.

7. Draper ES, Zeitlin J, Fenton AC, et al. Investigating the variations in survival rates for very preterm infants in 10 European regions: the MOSAIC birth cohort. *Arch Dis Child Fetal Neonatal Ed* 2009;94:F158-63.

8. Isayama T, Lee SK, Mori R, et al. Comparison of mortality and morbidity of very low birth weight infants between Canada and Japan. *Pediatrics* 2012;130:e957-65.

9. Stoll BJ, Hansen NI, Bell EF, et al. Neonatal outcomes of extremely preterm infants from the NICHD Neonatal Research Network. *Pediatrics* 2010;126:443-56.

10. Shah V, Warre R, Lee SK. Quality improvement initiatives in neonatal intensive care unit networks: achievements and challenges. *Acad Pediatr* 2013;13(Suppl):S75-83.

11. Lee SK, Aziz K, Singhal N, et al. Improving the quality of care for infants: a cluster randomized controlled trial. *CMAJ* 2009;181:469-76.

12. ViviWeb. EPIQ Web. Edmonton: Centre for Health Evidence; 2013. Available: https://meta.cche.net/viviweb/default.asp?site=EPIQ&lang=ENG (accessed 2013 Oct. 31). Login required to access content.

13. Shennan AT, Dunn MS, Ohlsson A, et al. Abnormal pulmonary outcomes in premature infants: prediction from oxygen requirement in the neonatal period. *Pediatrics* 1988;82:527-32.

14. Papile LA, Burstein J, Burstein R, et al. Incidence and evolution of subependymal and intraventricular hemorrhage: a study of infants with birth weights less than 1500 gm. *J Pediatr* 1978;92:529-34.

15. The International Classification of Retinopathy of Prematurity revisited. *Arch Ophthalmol* 2005;123:991-9.

16. Bell MJ, Ternberg JL, Feigin RD, et al. Neonatal necrotizing enterocolitis. Therapeutic decisions based upon clinical staging. *Ann Surg* 1978;187:1-7.

17. Garner JS, Jarvis WR, Emori TG, et al. CDC definitions for nosocomial infections, 1988. *Am J Infect Control* 1988;16:128-40.

18. CNN abstractors manual v2.1.1. Toronto: Canadian Neonatal Network; 2013. Available: www.canadianneonatalnetwork.org/portal/CNNHome/Publications.aspx (accessed 2013 Aug. 31).

19. Plsek PE. Quality improvement methods in clinical medicine. *Pediatrics* 1999;103(Suppl E):203-14.

20. Sullivan LM, Dukes KA, Losina E. Tutorial in biostatistics. An introduction to hierarchical linear modelling. *Stat Med* 1999;18:855-88.

21. Richardson DK, Corcoran JD, Escobar GJ, et al. SNAP-II and SNAPPE-II: simplified newborn illness severity and mortality risk scores. *J Pediatr* 2001;138:92-100.

22. Lee SK, Aziz K, Singhal N, et al.; Canadian Neonatal Network EPIQ-PHSI Study Group. The Evidence-Based Practice for Improving Quality (EPIQ) trial: reduction in nosocomial infection and broncholpumonary dysplasia incidence were sustained 2 years after the trial [poster abstract 1469.172]. Annual meeting of Pediatric Academic Societies; 2010 May 1–4; Vancouver.

23. Grimshaw JM, Thomas RE, MacLennan G, et al. Effectiveness and efficiency of guideline dissemination and implementation strategies. *Health Technol Assess* 2004;8:iii-iv,1-72.

24. Kaplan HC, Brady PW, Dritz MC, et al. The influence of context on quality improvement success in health care: a systematic review of the literature. *Milbank Q* 2010;88:500-59.

25. Horbar JD, Plsek PE, Leahy K. NIC/Q 2000: establishing habits for improvement in neonatal intensive care units. *Pediatrics* 2003;111:e397-410.

26. Johnston AM, Bullock CE, Graham JE, et al. Implementation and case-study results of potentially better practices for family-centered care: the family-centered care map. *Pediatrics* 2006;118(Suppl 2):S108-14.

27. Kuzma-O'Reilly B, Duenas ML, Greecher C, et al. Evaluation, development, and implementation of potentially better practices in neonatal intensive care nutrition. *Pediatrics* 2003;111:e461-70.

28. Payne NR, LaCorte M, Karna P, et al. Reduction of bronchopulmonary dysplasia after participation in the Breathsavers Group of the Vermont Oxford Network Neonatal Intensive Care Quality Improvement Collaborative. *Pediatrics* 2006;118(Suppl 2):S73-7.

29. Kilbride HW, Wirtschafter DD, Powers RJ, et al. Implementation of evidence-based potentially better practices to decrease nosocomial infections. *Pediatrics* 2003;111:e519-33.

30. Lee HC, Kurtin PS, Wight NE, et al. A quality improvement project to increase breast milk use in very low birth weight infants. *Pediatrics* 2012;130:e1679-87.

31. Lee HC, Lyndon A, Blumenfeld YJ, et al. Antenatal steroid administration for premature neonates in California. *Obstet Gynecol* 2011;117:603-9.

Affiliations: Department of Paediatrics (Lee, Shah), University of Toronto, Toronto, Ont.; Department of Pediatrics (Singhal), University of Calgary, Calgary, Alta.; Department of Pediatrics (Aziz), University of Alberta, Edmonton, Alta.; Department of Pediatrics (Synnes), University of British Columbia, Vancouver, BC; Department of Pediatrics (McMillan), Dalhousie University, Halifax, NS; Department of Pediatrics (Seshia), University of Manitoba, Winnipeg, Man.