Appendix 3 (as supplied by the authors): Details of statistical analysis

QI analyses

To track the impact of all QI initiatives we will monitor sites' progress by analyzing the common cause and special cause variations of process measures using upper and lower control limits (UCL and LCL) in a Shewhart chart (control chart or P chart) every three months and identifying positive changes.(21;43) When special cause variation is not identified, sites will assume that the process is "stable" and sites will implement the next QI strategy. In the event of a practice change with the potential for unintended consequences, the team will use the lower control limit (2 sigma, rather than 3 sigma) to abandon the practice. Sites' progress will also be analyzed using benchmarking. For each outcome, we will calculate adjusted rates and associated 95% confidence intervals and present them graphically using caterpillar plots to visually differentiate between EPIQ sites. We will use indirect standardization to estimate standardized ratios and adjust for multiple baseline characteristics and differences in important health care practices. The expected number of events will be computed as the sum of predicted probabilities from a Bayesian hierarchical generalized linear model (multilevel logistic regression or zero inflated negative binomial models, depending upon data distribution) of all individual patient data from all participating sites adjusted for patient and hospital level attributes. The Bayesian model will also allow us to stabilize the rate for small hospitals with few admissions. We will graphically display site standardized ratios using funnel plots with 95% prediction intervals to examine the variation between sites within the PTB Network.

Appendix to: Shah PS, McDonald SD, Barrett J, et al. The Canadian Preterm Birth Network: a study protocol for improving outcomes for preterm infants and their families. *CMAJ Open* 2018. DOI:10.9778/cmajo.20170128.

Comparisons and trends analysis

The PTB Network will collect data from the baseline year, 2018, until the last intervention year, 2020. We will compare infant characteristics over years, using a Chi-square test, ANOVA or Wilcoxon Rank Sum test, as appropriate. To test trends in infant characteristics; temporal trends; and variations in PTB rates, phenotypes, and subtypes (e.g., late, moderate, and extreme PTB; spontaneous and medically-indicated PTB) in Canada we will use a Cochran-Armitage trend test or linear regression as applicable. Overall outcomes will be compared using the Cochran-Mantel-Haenszel test to account for clustering within each site and Bayesian hierarchical generalized linear models adjusting for confounders, risk factors, and site characteristics. If applicable, we will perform piecewise hierarchical linear (over time) regression modeling with autoregressive covariance structure and a time series approach, adjusting for patient and site level factors, to assess the effectiveness of the interventions. Sub-analysis stratified by GA (< 26 weeks' and 26 – 29 weeks' GA groups) will be conducted.

Appendix to: Shah PS, McDonald SD, Barrett J, et al. The Canadian Preterm Birth Network: a study protocol for improving outcomes for preterm infants and their families. *CMAJ Open* 2018. DOI:10.9778/cmajo.20170128.