Neonatal Intensive Care Unit Collaboration to Decrease Hospital-Acquired Bloodstream Infections: From Comparative Performance Reports to Improvement Networks

Joseph Schulman, MD, MS[a,b,*], David D. Wirtschafter, MD[c], Paul Kurtin, MD[d]

KEYWORDS

- Quality improvement
- Performance evaluation • Risk adjustment
- Central line associated bloodstream infection
- Improvement networks • High reliability organization

Left to intuition, people have trouble accurately judging their relative performance.[1] They are prone to overestimate their achievements and capabilities compared with their peers—the Lake Wobegon Effect,[2] named for the fictional town of Lake Wobegon from the radio series *A Prairie Home Companion*, where "all the children are above average." Health care workers too, are subject to this bias,[3] particularly if their notion

[a] Division of Newborn Medicine, Department of Pediatrics, Weill Cornell Medical College of Cornell University, The New York Presbyterian Hospital, 525 East 68th Street, Box 106, NY 10065, USA
[b] Division of Outcomes & Effectiveness, Department of Public Health, Weill Medical College of Cornell University, The New York Presbyterian Hospital, 525 East 68th Street, Box 106, NY 10065, USA
[c] David D Wirtschafter, MD, Inc., 5523 Voletta Pl, Valley Village, CA 91607, USA
[d] Sadler Center for Quality, Rady Children's Hospital San Diego, 3020 Children's Way, San Diego, CA 92123, USA
* Corresponding author. Division of Newborn Medicine, Department of Pediatrics, Weill Medical College of Cornell University, The New York Presbyterian Hospital, 525 East 68th Street, Box 106, NY 10065.
E-mail address: jos2039@med.cornell.edu (J. Schulman).

Pediatr Clin N Am 56 (2009) 865–892
doi:10.1016/j.pcl.2009.06.001 **pediatric.theclinics.com**
0031-3955/09/$ – see front matter

of quality is based largely on good intentions. Learning to provide high-quality care may be considered to derive from understanding "The degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge."[4] This two-part article explores several performance improvement approaches motivated by this perspective and designed to reduce hospital-acquired bloodstream infections (BSIs) in neonatal intensive care units (NICUs). Readers may be reassured that the rather specific details that follow serve to illustrate concepts with broad applicability for pediatric quality improvement.

PART 1: COMPARING INSTITUTIONAL PERFORMANCE

As suggested in the title, the critical first step in quality improvement is making providers aware of their practices and patient outcomes. Because of the paucity of evidence supporting pediatric practice, many comparative performance reports are based on provider consensus about what is important to know and measure. After describing provider performance, the critical next step is deciding what to do with that information. Because most providers are not trained in the science and practice of quality improvement, quality improvement networks/learning collaboratives/communities of practice play an essential role in disseminating the knowledge, tools, and techniques of quality improvement and in helping to actually implement improvement plans.

Comparative Neonatal ICU Hospital-Acquired Bloodstream Infection Performance Reports: How are We Doing? Do We Have a Problem?

Imagine you work at the hospital represented by the bold dot in **Fig. 1**. The display plots NICU nosocomial sepsis infection (NI) outcomes for several hundred hospitals participating in the Vermont Oxford Network.[5] Each dot represents the difference between the number of positive blood cultures in infants after day 3 of life reported by an individual hospital and the number that would be expected at that hospital after accounting for specific differences in patient characteristics—case mix—from hospital to hospital. Additionally, the dots reflect so-called "shrunken" estimates: a method that accounts for the assumption that one infant's infection is not independent of another's, and adjusts for individual hospital estimates considered relatively imprecise by moving such values closer to the mean value for all hospitals. So, how is your facility doing with these hospital-acquired infections (HAIs)? Do you have a problem with this aspect of care?

Displays as this one may appear pretty straightforward at first glance; but only at first glance. Each dot encodes the result of a complex analysis that rests upon crucial assumptions and important limitations. Tempting as it may be immediately to articulate the "answer" provided by a given dot, for several reasons, an impulsive determination is fraught with risk. First, the answer represented by each dot depends on the question posed—the hypothesis tested. Second, meaningful and fair comparison of various institutions' results depends on accounting for baseline patient characteristics that importantly influence the outcome of interest but that are independent of quality of care. Third, valid results depend on an analysis that conforms to methodological assumptions. And fourth, correctly drawn inference integrates these three considerations with explicit assessment of residual chance, bias, or confounding.[6] Let's explore some of these ideas further before we decide what the bold dot tells us about clinical performance.

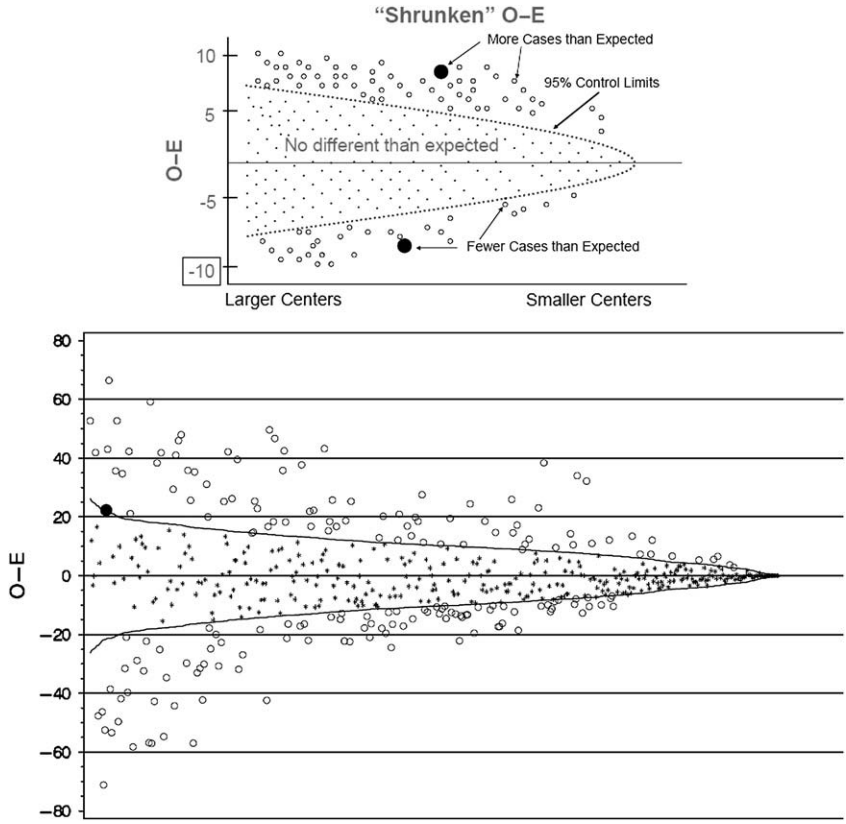

Fig.1. (*Top*) Interpretation key for lower image—observed minus expected cases adjusted for baseline characteristics and random effects. Each dot represents an individual hospital. Horizontal axis depicts absolute value of 95% control limit, considered a proxy for number of patients at a hospital (*Courtesy of* Vermont Oxford Network, Burlington, VT; with permission.) (*Lower*) Instantiates this configuration, reporting on nosocomial infection; the heavy black dot is the center of interest. The shape of the superimposed curve, marking the range of variation attributable to chance, that is, ±95% control limits—essentially 2 standard deviations for process-related data, suggests the name: funnel plot. This feature displays precision of the estimate as a function of hospital volume. Dots within the control limits are statistically indistinguishable, which means it makes no sense to rank-order those hospitals. (*From* Schulman J, Spiegelhalter DJ, Parry G. How to interpret your dot: decoding the message of clinical performance indicators. [State of the Art Article.] J Perinatol 2008;28:588–96; with permission).

Are the comparisons fair?

At the outset, it is worth discriminating at least two different and important uses for quality improvement data:

(1) Data for improvement, which do not support operationally specific inferences (concrete ideas about what to do differently) to the extent we might like, but the data nonetheless can inspire hypotheses about how to improve; and then those hypotheses must be tested.

(2) Data for accountability/external comparisons. These data are made public. They ought to support operationally specific inference and be as reliable and valid, as possible; ie, providers must believe the data.

Although the authors consider perfect risk adjustment in the NICU an impractical ideal, it is reasonable to compare performance as a means to seek opportunities for internal improvement…not external accountability.

It is crucial to appreciate that accounting for specified baseline differences among individual group members, for example, patients in various hospitals, doesn't necessarily mean that groups—hospitals—are being compared fairly. Some exposure/outcome relationships may be distorted when the outcome is also substantially affected by *un*accounted differences among individual group members. More concretely, the relationship between receiving care in a particular facility (primary exposure) and risk of developing an HAI (outcome) can change when other baseline differences—risk factors, secondary exposures—are included in the analysis. To illustrate, let's compare the incidence of HAI in imaginary NICU A and NICU B (**Table 1**).

Table 1
Incidence of hospital-acquired infection in imaginary NICU A and NICU B

All Subjects	**NICU A**	**NICU B**	**Total**
HAI	45	30	75
Infection-free	455	470	925
Total	500	500	1000
Risk	.09	.06	.075
—	Point estimate		—
Risk difference	.03		—
Risk ratio	**1.5**		—
MALES	**NICU A**	**NICU B**	**Total**
HAI	5	20	25
Infection-free	95	380	475
Total	100	400	500
Risk	.05	.05	.05
—	Point estimate		—
Risk difference	0		—
Risk ratio	**1**		—
FEMALES	**NICU A**	**NICU B**	**Total**
HAI	40	10	50
Infection-free	360	90	450
Total	400	100	500
Risk	.1	.1	.1
—	Point estimate		—
Risk difference	0		—
Risk ratio	**1**		—

The risk ratio (RR) for developing a hospital-acquired infection (HAI) in neonatal ICU (NICU) A compared with NICU B is 1.5 when all subjects are considered, but RR = 1 when stratified by gender, because gender is associated with site of care (NICU A versus NICU B) and with outcome.

The risk ratio (RR) for developing an HAI in NICU A compared with NICU B is 1.5 when all subjects are considered: ie, patients in NICU A are at 50% increased risk of developing an HAI compared with patients in NICU B. However, when only males or only females are considered, the RR is 1 (no risk difference). How can an apparently strong exposure effect disappear when we stratify by gender?

At the heart of this question is a phenomenon called Simpson's Paradox. The RR of 1.5 among all subjects disappears when analysis is stratified by gender because gender is associated with both primary exposure—the NICU in which a patient receives care, and with outcome—HAI. No real association exists between primary exposure and outcome. Rather, the association is between *secondary* exposure—*gender*—and outcome. The apparent relationship reflects unequal distribution of males and females in NICU A and NICU B. The point to underscore is that each analysis answered a different question:

- "Among the **entire** study group, what is the association between primary exposure and outcome?"
- "Among the **males** in the study group, what is the association between primary exposure and outcome?"
- "Among the **females** in the study group, what is the association between primary exposure and outcome?"

So, if apparent exposure/outcome relationships can change as additional factors are accounted for, *how does one gain confidence that an exposure/outcome relationship is real—that the important factors have indeed been accounted for?* The key idea is that patients at different hospitals may differ in various ways and some of those differences might substantially explain observed variation in outcomes among hospitals. In general, meaningful comparisons ultimately entail comparing like with like. So metaphorically speaking, the problem of comparing hospital performance centers on this question: "If asked to evaluate several bowls of fruit salad, how does one transform the diversity into a comparison of apples with apples?"

Statisticians try to achieve this via so-called risk adjustment methods. Sometimes the magic works very well and other times we are left still comparing apples with oranges, pears, and so on. Several measures exist to quantify how well a risk adjustment model has achieved the desired result.[7] Clinicians should understand at least two of these. (We emphasize that critical review is not intended to create an "excuse" for taking no action. An analytical review may be critical yet constructive by enumerating potential limitations along with speculating about the direction and magnitude of potential bias.[8] For example, in light of the identified limitations perhaps the findings represent an overestimate (or underestimate) of the effect of site of care on outcome. On the other hand, if such review suggests that the *direction* of the relationship could be different, then it may indeed be appropriate to suspend action until better data are available.)

Evaluating a risk adjustment model

One measure of model performance is discriminatory power: how well a model, on the basis of specified baseline characteristics, discriminates subjects who will experience the outcome from those who will not. Discriminatory power may be quantified as the area under a plot of model sensitivity versus 1-specificity: the receiver operating characteristic (ROC) curve (**Fig. 2**).[9] Values range from 0 to 1: 0.5 = random outcome assignment (coin toss); 1 = perfect discrimination. The *c*-statistic, using the same range of values, quantifies discriminatory power by evaluating all pairs of subjects, one of whom experienced the outcome and one of whom did not; and evaluates the

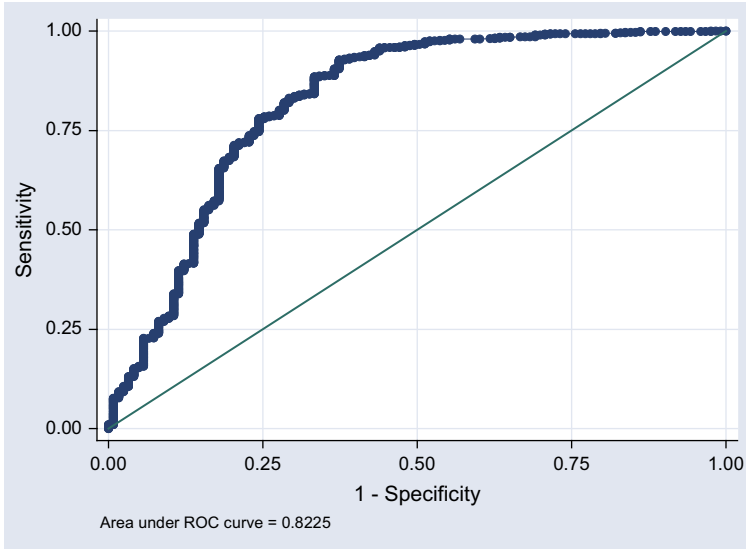

Fig. 2. Illustrative ROC curve. (*From* Schulman J. Managing your patients' data in the neonatal and pediatric ICU: an introduction to databases and statistical analysis. Oxford, UK: Blackwell; 2006; with permission.)

proportion of pairs for whom the model correctly predicts higher probability of experiencing the outcome for the subject who actually experienced the outcome.

A second model performance measure is explanatory power: the extent to which the model can explain the observed variation in outcome among subjects. Models use observed values to produce an equation summarizing the risk factor/outcome relationship. This equation predicts outcome values expected for a given profile of risk factor values; thereby describing a line. In reality, however, *observed* values don't fall precisely on the prediction line.

More specifically, models predicting a dichotomous outcome—yes/no; lived/died; infection/infection-free—assign a predicted probability of outcome occurrence to each individual subject based on that subject's risk factor values. Now, probabilities range from 0 to 1, but patients either experience the outcome or do not, so observed values are all either 0 or 1 (**Fig. 3**). Thus, observations off the line described by the predictive equation contain information that the model could not use. In general, the extent to which a model could use the observed information is measured by a quantity called R^2, or explanatory power. Theoretical R^2 values range between 0 (no association between predicted and observed values) and 1 (perfect association). The extent to which a model could *not* use information contained in the observations is measured by $1 - R^2$, also called unexplained variance. The latter represents the contribution of unmeasured factors to the observed variance. So if $R^2 = 0.16$, the model fails to explain 84% of observed variance. For dichotomous outcomes, R^2 reflects the difference between average predicted outcome probability among those who experienced the outcome and those who did not,[8] and commonly is reported as pseudo R^2. Models that account well for determinants of an infrequently occurring outcome nonetheless will have a low pseudo R^2. For example, when observed infection rate is 2% to 4%, pseudo R^2 is usually less than 0.25, because the difference between the predicted and observed values commonly is large. (Robust severity adjustment systems do

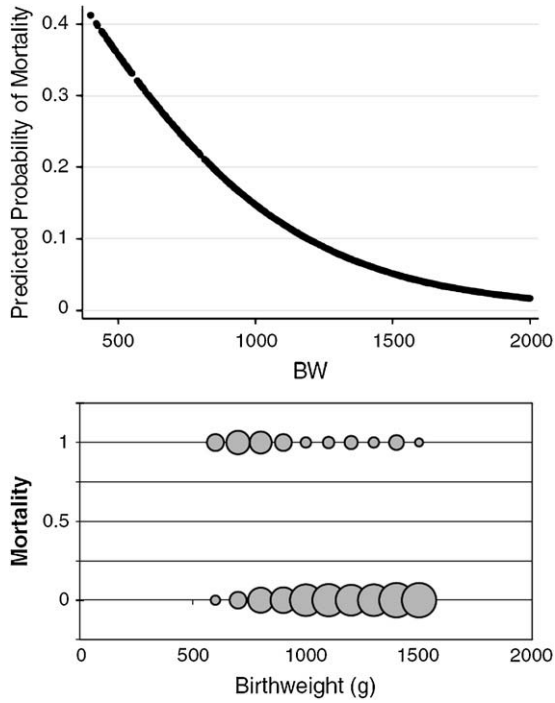

Fig. 3. Left, predicted survival probabilities as a function of birth weight (BW); right, observed outcomes by hospital, circle diameter reflects hospital volume. (*From* Schulman J, Spiegelhalter DJ, Parry G. How to interpret your dot: decoding the message of clinical performance indicators. [State of the Art Article.] J Perinatol 2008;28:588–96; with permission.)

not exist for much of pediatrics, so we need to accept imperfect risk adjustment and move forward because ultimately improvement means comparing self to self over time; if the patient population in a given NICU stays fairly similar over time then that comparison is reasonable. However, we do need agreement on operational definitions such as how to diagnose a central line associated bloodstream infection (CLABSI) or whether ''clinical sepsis'' is to be included in the definition. Voltaire made the essential point: ''the best is the enemy of the good.'')[7] One upshot of this is that detecting a change in occurrence rate for infrequent events requires both a large sample size and substantial change in risk. For example, if baseline occurrence rate is 1%, 747 subjects with the risk factor of interest and 747 subjects without that factor are needed to detect a threefold increased risk.[10]

Explained and unexplained variance are crucial ideas because some organizations consider unexplained variance a surrogate for ''quality.'' That is, outcome differences remaining among hospitals after risk adjustment are considered to reflect care at the respective hospital. Doing so rests on two important assumptions:

1. No other factors that are independent of quality of care operate to substantially influence the relationship between site of care and outcome. That is, one is unable to think of other baseline patient characteristics that might have a substantial effect on the outcome. If otherwise, those other factors should be included in the model to

provide reassurance that exposure/outcome relationships nevertheless remain stable.
2. An unimportant proportion of the remaining variance is due to chance. The smaller the explanatory power and the larger the unexplained variance, the larger the potential contributions of irrelevant information and random error to the model; and therefore the more tenuous the inference about quality.

"Shrunken" estimates: multilevel risk adjustment models

One more type of special accounting merits further explanation to help decipher the message of the bold dot in **Fig. 1**: so-called ''shrunken'' estimates. Traditional risk adjustment methods assume that observations are independent of each other; for example, that a hospital-acquired infection in one patient has nothing to do with such an event in another patient in the same facility. However, what happens to one patient may not be independent of what happens to another patient in the same facility. Factors that predispose one patient to an outcome may also operate to predispose other patients to that outcome. Outcomes may be correlated: systematic differences unexplained by specified predictors may exist among hospitals. (Quality improvement is often based on understanding systemic factors as a root cause to differential performance, eg, hand hygiene rates may impact many outcomes and may be performed at very different rates by different members of the staff at different units. Other examples include presence of a PICC team; average length of time the nursing staff has worked in a given unit; having a dedicated catheter cart; who has responsibility for PICCs, housestaff versus nurse practitioners; parents' hand hygiene compliance rates.) One must therefore account for the varying roles of individual-level and group-level factors in jointly determining outcomes. This may be done via multilevel modeling, which yields ''shrunken'' outcome estimates. The predicted result for an individual hospital tends to ''shrink,'' away from the hospital's observed value toward the value of the summary measure for all hospitals. How much a hospital's predicted value shrinks depends on three criteria for assessing the reliability of a particular hospital's observed result:

1. *Within-group variation*: How much do individual patients' values differ from the hospital's average outcome?
2. *Between-group variation*: To what extent does each hospital differ from the average for all hospitals?
3. *Patient volume*: Low-volume hospitals are considered a relatively less reliable data source.

Reliability here relates to the extent an observed value may have been affected by chance events and not the system of care. When an individual hospital's observations seem less reliable, the model assigns less weight to that hospital's observed outcomes and more weight to the overall experience of all hospitals.

So, how are we doing?

We are now prepared to extract the message of **Fig. 1**. The particular question the display answers—the hypothesis that was tested statistically—is not explicitly articulated in the report document provided to the institution represented by the bold dot. Instead, as is common with performance reports, one is simply provided with an implicit ''answer,'' in this case, ''observed minus expected nosocomial infections.'' (The model of improvement is built upon the necessity for a clearly defined AIM because from that aim will come the measures that are used to create the appropriate

baseline as well as the means to assess progress.) Several questions are possible, including:[7]

1. ''Which hospitals' results are not compatible with the overall average?''
2. ''Which hospitals are in the top or bottom half?''
3. ''Does a particular hospital come from the same distribution as all the others; is this hospital an extraordinary performer?''

The question actually posed determines the particular conformation of the funnel-shaped curve in **Fig. 1**. The report document specifies risk factors used in the risk adjustment model, but not the model's discriminatory or explanatory power; cautioning, however, that risk adjustment methods are imperfect. Nevertheless, thoughtful interpretation of risk-adjusted results depends on some factual understanding of just *how imperfect* is the comparison. ''All models are wrong; the practical question is how wrong do they have to be to not be useful'' (George Box, quoted in Schulman).[11]

Mindful of the foregoing caveats we can say this: lying as it does on the superimposed funnel-shaped curve, our dot in **Fig. 1** represents a borderline ''outlier.'' The funnel curve marks the range of variation attributable to chance, ±95% control limits. Variation among dots within the control limits is more likely attributable to chance than to inherent differences in systems of care.

So, what does it mean to be an outlier? Only that if the model has done a good job transforming the evaluative task to a comparison of apples with apples, then we can be confident the dot differs from the overall average, the model expectation. However, shrunken estimates ''*assume* that systematic differences exist among the hospitals; that ALL hospitals differ from the average, so confirming this in a particular case is hardly noteworthy!''[7] It is important to appreciate that larger hospitals with slightly excess risk are prone to be classified as outliers—they probably differ from the overall average, the model expectation—because the statistical method recalibrates low volume hospitals' performance so that they approximate the overall average while high volume hospitals' values change little. Thus nearly all hospitals that care for a sufficiently high patient volume will lie outside the control limits.[7]

Neonatal ICU Hospital-Acquired Bloodstream Infection Improvement Networks

The previously enumerated questions that might motivate **Fig. 1** center the focus of data-driven quality improvement efforts on individual data points describing institutional performance. Question number 3 is potentially most helpful. Outliers revealed by this hypothesis test might provide insight to structural and process-related variables whose relationships to outcomes could then be tested among the entire group.

An alternate focus with great improvement potential centers on the observed distribution of the data points: Can we favorably shift average performance of the group overall and narrow the variation among individuals? Part 2 of this article describes improvement initiatives by neonatologists in New York and California illustrating some of the possible strategies.

PART 2: COLLABORATING TO IMPROVE MANY INSTITUTIONS' PERFORMANCE

Neonatal ICU Hospital-Acquired Bloodstream Infection Improvement Networks

Part 1 of this article provided key considerations for fairly comparing individual institutional performance. Useful evaluation was shown to derive from testing a hypothesis resembling ''Does a particular hospital come from the same distribution as all the others; ie, is this hospital an extraordinary performer?'' In this second part, the focus

shifts from individual data points (hospitals) to the observed distribution of the data points; and the corresponding hypothesis becomes ''Can we favorably shift average performance of the overall group and narrow the variation among individuals?'' Improvement initiatives by neonatologists in New York and California illustrate possible strategies.

New York

Background, methods

Early in 2007 all 19 regional perinatal centers (RPCs) of New York State (NYS), supported by the New York State Department of Health (NYSDOH) and the Association of Regional Perinatal Provider Networks, collaborated to decrease central line associated bloodstream infection (CLABSI) rates in their neonatal intensive care units (NICUs).[12] RPCs are State-designated to serve ''…the most acutely sick or at-risk pregnant women and newborns… coordinate … transfers of high-risk patients from their affiliate hospitals to the RPC, and are responsible for support, education, consultation and improvements in the quality of care in the affiliate hospitals within their region.''[13] Representation was multidisciplinary: front-line physicians, nurse practitioners, nurses, and infection control professionals.

Despite mandated infection reporting, until this initiative began individual RPCs were unaware of how their results compared with others'. RPCs only knew their referring hospitals' performance, not other RPCs'. The extent to which centers applied best practices to prevent CLABSIs was unclear. Centers suspected that CLABSI rates and pertinent practices varied substantially among RPCs. Centers recognized that they needed an appropriate, credible standard by which to evaluate RPC performance and motivate improvement.

During the collaborative's first year, RPCs shared self-identified NICU bloodstream infection rates and agreed to set a standard of care via a bundle of practices – percutaneously inserted central catheters (PICCs) were specified because this line type accounted for the bulk of central line days. NYS Public Health Law Article 28 § 2819 required that all NYS hospitals join the National Healthcare Safety Network (NHSN),[14] which made CLABSI rate available statewide. Every individual participant had to join either a Research Panel (RP) or a Quality Improvement Panel (QIP); each established defined individual responsibilities. The RP selected performance indicators; reviewed data completeness, integrity, and analysis; and identified candidate host institutions for participants to visit and learn about potential best practices. The QIP reviewed pertinent peer-reviewed literature; coordinated the visit to a potential best performing NICU (with respect to the specified indicator; ideally, one could look for the best performance on a number of indicators, as different places often do different things well); and spearheaded creation of a care bundle. Controversy arising from collaborative activities was adjudicated by both panels. Several dozen individuals visited a NICU with low indicator rate to learn about potential best practices. Many had never before scrutinized specific care practices so carefully. Visitors observed morning rounds and hand hygiene practices; hosts described care practices for catheter insertion, maintenance, and removal; formal debriefing consolidated learning. Key ''take-away'' messages included the following:

- Facilitate adherence to hand hygiene practices: place hand sanitizer dispensers at each bedside.
- Use two staff members for central line placement.
- Change catheter tubing by nurse pairs using a ''buddy system,'' according to a written policy.

- Consider line removal each day during patient rounds.
- Create a culture wherein the NICU staff, a physician ''champion'' and the infection control professionals consider themselves a single team working to prevent CLABSI.

Elements of the evidence-based practice bundle to prevent CLABSIs were required to derive from level-1 evidence.[15] Pertinent literature identified by the QIP was distributed to all RPCs for review.

Results

The 19 RPC NICUs shared 56,911 central line days of observation; **Fig. 4** shows birth weight (BW)-stratified and aggregate institutional performance (details available in Schulman and colleagues[12]) Each participating RPC knew which letter identified each institution.

Informed by Part 1 of this article, critical readers might now ask: ''If the lower display in **Fig. 4** is the 'answer,' what's the question?'' The question that indeed motivated the analysis was: ''During 2007, what was the overall CLABSI rate at each NICU, and assuming no changes in circumstances—for example, stable patient demographics, stable processes of care, stable risk factor incidence—what is the underlying CLABSI risk for new patients at a particular NICU?'' The lower display does not answer: ''Do rates sufficiently account for differences in baseline patient characteristics among NICUs?'' Nor does it answer: ''How much must values differ to 'mean something'—Is variation among centers more likely due to chance than to inherent differences in their systems of care?'' The actual rates depicted by each dot in **Fig. 4** are called point estimates; they are accompanied by 95% confidence intervals (CIs). Even with constant underlying risk, the proportion of neonates who will develop a CLABSI at each NICU varies every evaluation interval. CIs help to answer: ''What is the true CLABSI rate for each NICU?'' in the sense of establishing a *range* of underlying risk values compatible with observed outcomes. So, the overall CLABSI rate for NICU B during 2007 was approximately 12.7/1000 central line days. However, repeating the observations over a very long period, all else unchanged, the overall CLABSI rate for NICU B should lie between 4.6 and 27.6/1000 central line days. The computed interval width is an inverse function of patient volume. Larger samples will more accurately reflect the true population characteristics, ie, the individual estimate is more precise; CI narrows (compare NICU B with NICU J).[7]

Within each BW category CLABSI rates varied widely. Combining all BW strata, individual center rates ranged from 2.6 to 15.1 per 1000 central line-days, representing a nearly sixfold variation across the participating NICUs. For infants of BW less than 751 g rates ranged from 0 to 19.0 per 1000 central line days: nearly 20-fold variation across NYS's regional referral NICUs. It is unclear how much between-center variation is a result of chance (note the extensive overlap of CIs in **Fig. 4**).

The central line care bundle adopted by unanimous agreement is shown in **Box 1**.

Discussion

Health care has entered an era of mandatory statewide reporting of CLABSI rates for ICU patients.[30,31] Neonatologists at New York State's regional referral centers reframed mandated HAI reporting as an opportunity: to promote accountability and cooperation among providers and to develop a strategy to prevent CLABSIs. The strategy of a statewide collaboration is innovative for several reasons. First, because all RPCs participate, potential sampling error or bias is avoided and population-based rates can be determined. Second, because evaluated providers are directly involved in data collection, review, analysis, and interpretation, providers believe performance

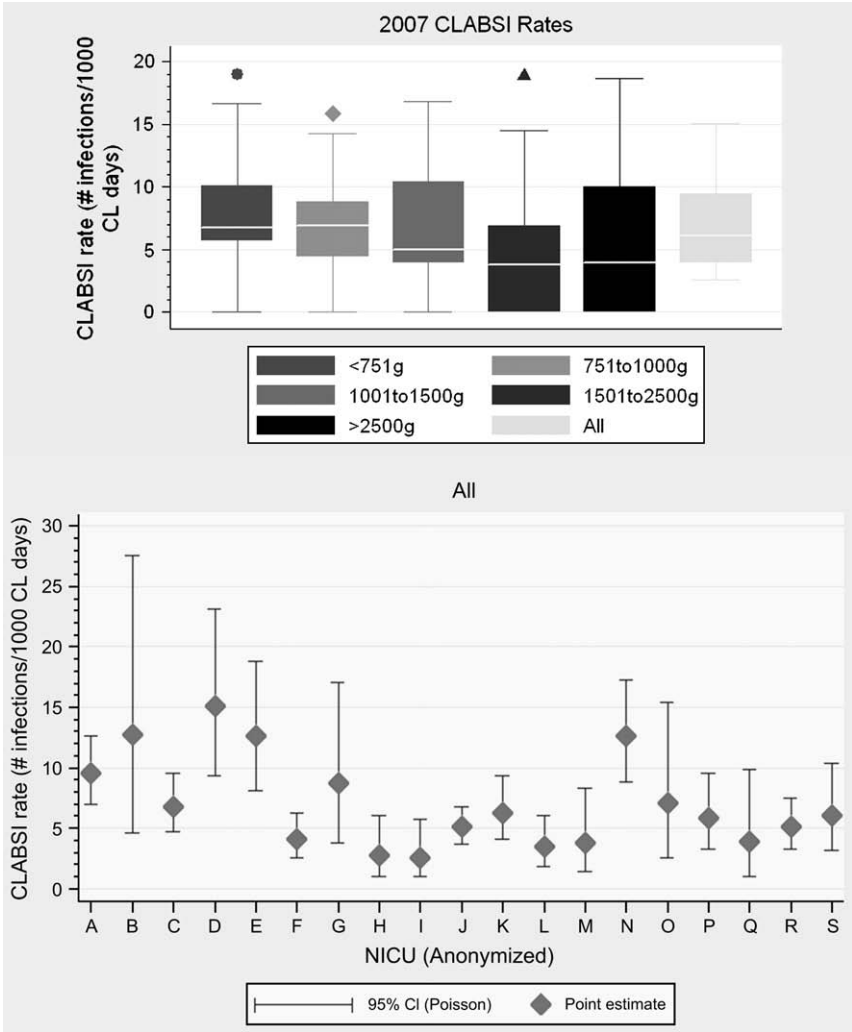

Fig. 4. (*Upper*) Range of BW-stratified CLABSI rates for participating RPCs during the calendar year 2007. Each box bounds values between the first and third quartiles, with the horizontal line within a box denoting the median value. The interquartile range (IQR) is computed by subtracting the value of the first quartile from the value of the third quartile. Values that exceed the third quartile by more than 1.5 times the IQR are identified as dots. (*Lower*) Between-RPC variation in CLABSI rate. CIs assume Poisson distribution, 95% interval. (*From* Schulman J, et al. Development of a Statewide Collaborative to decrease NICU central line associated bloodstream infections. J Perinatol, in press; with permission.)

results. If they do not, they are empowered to identify and correct their data thereby improving data accuracy. Third, the NYSDOH Office of HAI conducted site visits to assess sites' use of CLABSI case definitions and adjudicated cases.[32] Fourth, the network intends to help members improve by monitoring bundle element application and effectiveness.

Box 1
New York State Regional Perinatal Care Centers: PICC bundle

This tool specifies vital care elements to be consistently applied by credentialed practitioners

PICC Insertion

- Establish a central line kit or cart to consolidate all items necessary for the procedure (IB).[16]
- Perform hand hygiene with hospital-approved alcohol-based product or antiseptic-containing soap before and after palpating insertion sites and before and after inserting central line (IA).[17–19]
- Use maximal barrier precautions (including sterile gown, sterile gloves, surgical mask, hat, and large sterile drape) (IA).[17,20]
- Disinfect skin with appropriate antiseptic (eg, 2% chlorhexidine, 70% alcohol) before catheter insertion (IA).[17,21,22]
- Use either a sterile transparent semipermeable dressing or sterile gauze to cover the insertion site (IA).[23–25]

PICC Maintenance

- Perform hand hygiene with hospital approved alcohol-based product or antiseptic-containing soap before and after accessing a catheter or changing the dressing (IA).[17–19]
- Evaluate the catheter insertion site daily for signs of infection and to assess dressing integrity (IB).
- At a minimum, if the dressing is damp, soiled, or loose change it aseptically (IB) and disinfect the skin around the insertion site with an appropriate antiseptic (IA).[21,24,26,27]
- Develop and use standardized intravenous tubing set-up and changes (IB).[28]
- Maintain aseptic technique when changing IV tubing and when entering the catheter including "scrub the hub" (IA).[24,27,29]
- Daily review of catheter necessity with prompt removal when no longer essential (IB).[16,17]

Category IA. Strongly recommended and strongly supported by well-designed experimental, clinical, or epidemiologic studies.
Category IB. Strongly recommended and supported by some studies and strong theoretical rationale.
Centers for Disease Control and Prevention[15]

Note that *no center was without CLABSI*. Thus, the extent to which these events may be prevented, even in the best performing centers, remains uncertain.[33] Nonetheless, in adult and pediatric ICU populations, implementation of bundle strategies and checklists has been associated with substantially decreased CLABSI rates,[17,34] and the efforts of neonatologists in California, described later in this article, support the association for NICU populations. Thus, it is sensible for New York State neonatologists to try to shift average performance and narrow the variation among NICUs.

Although the bundle in **Box 1** derives from high-quality evidence, monitoring of its effectiveness is essential. The best practices collected in the bundle typically emerge by studying an individual intervention in a controlled setting (preferably a randomized controlled trial). However, in actual clinical care several best practices may be applied in combination; a procedure that assumes favorable interactions among the individual interventions. To further complicate the situation, best practices may yield varying results that depend on the myriad ways actual care settings and patients differ from the controlled circumstances operating when effectiveness was initially established.

(These are critical points because it is essential to deal with the reality of daily clinical practice – in which many changes are likely to occur simultaneously. [Unlike the randomized control trial, real-life practice is very hard to tightly control.] Organizations working to improve care will often do several interventions at once.)[35,36]

This rationale for monitoring bundle effectiveness also begs a question: Might asking whether an intervention works or not oversimplify the situation? It may be more illuminating to ask: ''For *whom* might it work, and *under what circumstances*?''[35,37] Delivery of health care is a social phenomenon. As such, it may be wishful thinking to suppose that study and control groups only differ in the intervention of interest, and that no other aspects of care vary due to human factors. Local details about ''how'' something works and about the ''what'' of contexts may be crucial.[37] For example, aggregated analysis of a multicenter trial suggesting that an intervention does not ''work,'' ie, the intervention has no effect, might obscure the reality that the intervention is beneficial for some types of patients in particular settings and harmful for other types of patients or settings. The quality improvement approach of neonatologists in California incorporates this idea.

California

Network overview

California's regional NICUs provide care to the majority of complex medical and surgical neonatal patients in the state. In 2006, California Children's Services (CCS) joined in a unique collaborative relationship with the California Children's Hospital Association (CCHA) to develop and implement a statewide initiative intended to reduce CLABSIs in these NICUs.

The project leadership team included CCS leadership (Marian Dalsey, MD, MPH, Kathy Chance, MD, and Hallie Morrow MD, MS), quality improvement expertise (Paul Kurtin, MD, Michael Seid, PhD, and Thomas Huber, M.) and neonatal medical/nursing expertise (David D. Wirtschafter, MD, and Janet S. Pettit, MSN, NNP-BC). In 2006, 13 members of CCHA joined the project (September 2006 to December 2007), and the remaining nine regional NICUs joined in January 2008. The 22 collaborating NICUs included all eight Children's Hospitals in the state, all five University of California Hospitals, and nine large regional centers, including those for three multisite hospital systems. NICU size ranged from 23 to 84 beds and patient days in 2007 ranged from 7665 to 29,565. In 2007, these 22 units provided 285,430 NICU patient care days. The original 13 NICUs had 59,182 line days in 2006 and 73,077 in 2007. The expanded cohort will report an estimated 87,500 line days in 2008.

Members worked collaboratively using three face-to face meetings to initially address project goals and methods for performing QI initiatives and later to share accomplishments, challenges and means to overcome them and form consensus around new approaches to reduce CLABSIs. Biweekly phone conferences served to update members of the collaborative on the activities of each individual site, identify potential interventions, develop consensus about data definitions and collection procedures, exchange knowledge about ''best practices,'' and exchange ideas on ''promising practices.'' The project's leaders provided one on-site consultation visit during 2007 to the original 13 members. These visits provided tailored consultations aimed at helping each site identify and overcome barriers to implementing necessary changes in practice. Annual project meetings included all participants and focused on lessons learned, opportunities for continued improvement, and project assessment, followed by a presentation to and celebration with the member's CEO's and CCHA's senior leadership.

Performance reports

Because we wanted to understand the performance of the collaborative as it was changing and to know whether real change was occurring, we used statistical process control (SPC) methods. Although traditional inferential statistics are commonly used to examine differences between samples from different sites or collaborative performance at different times, it requires aggregation of measures over time and thus can slow the pace of change. SPC is a branch of statistics that uses time series analysis and graphical presentation of data. Traditional inferential statistics rests on the assumption that sampled data are drawn from a stable population, whereas CLABSI rates reflect a reiterative production process whose results vary over time: ie, a NICU care system. SPC is designed to accurately characterize such variation over time—to discriminate signal from noise, how much of a change must occur for it to ''mean something''; and to indicate whether the system performs predictably or not. NICU A may have the same annual infection rate as NICU B but NICU A may operate within a substantially narrower performance range than NICU B, so that one may reliably predict the coming month's results in NICU A, but not in NICU B.[38] Researchers and practitioners often find SPC data displays very useful for evaluating performance and guiding change.[39]

We created performance reports (CLABSIs per 1000 central line days) for each participating site that highlighted the large degree of variation in rates among the participating sites and compared their results with the national NHSN benchmarks.[40] The reports were primarily used to (1) create a baseline of group performance, (2) create a baseline for each site to compare itself to its own performance over time, and (3) help make the case that changes in practice are needed to both reduce variation within and between sites and to improve performance compared with national benchmarks.

CLABSI rates during the first year of the project fell 25% from 4.32 to 3.22 infections per 1000 catheter days comparing the baseline 2006 period to the whole of 2007. This represented 75 fewer infections annually in a population exposed to 73,000 line days. **Fig. 5** shows a CLABSI rate control chart for the very low birth weight (VLBW) babies. There is a downward shift in the CLABSI rate from the baseline through the intervention period. Moreover, while there is some increasing scatter during the follow-up period, this is not a strong enough shift to suggest a rising CLABSI rate. During this post-intervention period (project hiatus: July to December 2007), there were no group phone conferences and no communication regarding individual performance rates occurred.

The collaborative's CLABSI rate was 2.36/1000 line days for January to September 2008. It is not possible to directly compare this rate with the rates obtained in the previous year because of a change in the stringency of the NHSN CLABSI definition implemented on January 1, 2008.[41] The new one calls for two rather than one positive blood cultures when diagnosing a catheter-associated infection from a recognized contaminant. Since coagulase-negative staphylococci (CONS) may constitute up to one half of the organisms associated with neonatal infection,[42] this criteria change has the potential to confound neonatal HAI trend analysis. Anecdotally, the California collaborative members report that perhaps one third of their 2007 CONS infections were not diagnosed with two blood cultures, and, even though treated, they would not have been designated as CLABSI events in 2008. A similar effect was estimated in New York State (Stricof RL, December 11, 2008, personal communication).

Identifying the right ways to do the right things

The project generated many different threads of discourse and discovery, both because there exists no single vascular access ''bundle'' for neonates and because

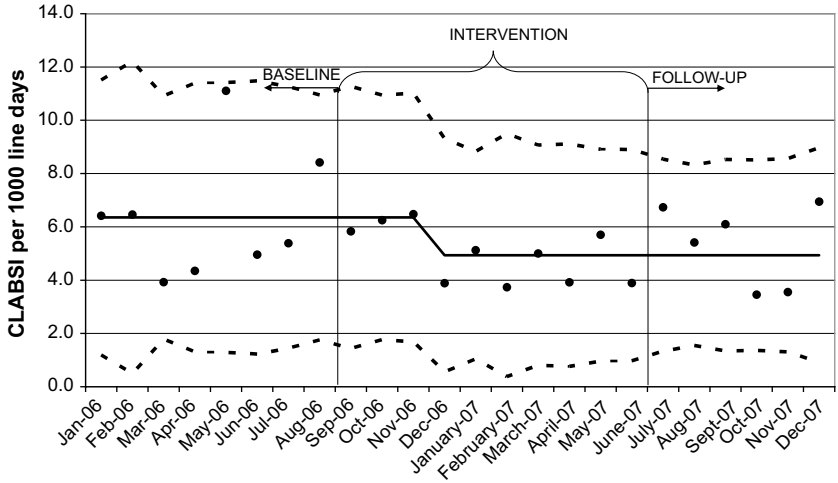

Fig. 5. CCHA-CCS NICU Collaborative: CLABSI rates among infants with birth weights <1500 g 2006–2007. Mean (*solid line*) and 99% confidence intervals (*dashed lines*) calculated using the method described by McCarty et al for binomially distributed data.

there are many technical and material challenges in meeting the needs of these patients. Thus, substantial portions of the collaborative's work consisted of developing and assessing ideas for use in the NICU. Some of the specific technical challenges considered were the following: whether, or with what restrictions, to use chlorhexidine in skin antisepsis; availability and workability of closed system vascular access devices for umbilical and peripheral vascular lines; designs of medication administration lines to lessen their exposure to bedside contaminants and to minimize the fluid load required to maintain their patency; standards on how to obtain blood cultures (considering both volume and site(s)); and revising the NSHN CLABSI criteria as they pertain to temperature assessment to reflect better the variability of neonatal temperatures and how they are measured. All of these issues, critical to the diagnosis and prevention of CLABSI, were unspecified at the beginning of the project, with no well defined and tested best practices available to be cited from the literature (**Table 2**).

Connecting upstream process steps with downstream outcomes

One of the key work products of this collaborative is a rich dataset describing the different paths taken by the individual centers in achieving their QI goals. This was possible because of our initial work to define prevention processes in greater and more practical detail than at the aggregated bundle level. The continuous real-time logging of member's work enables us to describe the progression of project activities.

The collaborative also developed a comprehensive means for assessing and recording the investigation of a positive blood culture report (**Fig. 6**). We derived its many suggested assessments from specific case reports, either published or suggested by our member's experience. Our member's anecdotal evaluation of this process indicates that a literature-based forensic approach stimulates a more far-reaching survey for both usual and unusual "suspect" practices potentially contributing to a CLABSI. Refusing to simplify explanations for unexpected events, as described above, is an example of the reflective practices of high-reliability organizations (HRO).[43]

The overarching challenge: promoting high reliability characteristics

We sought to identify and deploy feasible practices that could enhance these NICUs' adoption of ''HRO characteristics.'' For example:

- We deployed previously available observation tools or created de novo tools for assessing the adequacy of vascular line set-ups and entry, the adequacy of hand hygiene, and the real time investigation of all CLABSIs. When our members' audits indicated performance shortfalls, for instance, hand hygiene rates falling below 90%, they responded by initiating reinforcement and re-education of desired techniques. Recently, some of our members have merged the need for on-going audit of line set-up processes with the leadership need to ensure daily contact between managers and nurses.[44] In a process they describe as ''walk the line,'' they have asked nursing shift managers, aka ''charge'' nurses, to walk by each neonate's bed and review the line set-up and briefly interact with the bedside nurse. This practice also helps to enhance ''situational awareness'': being mindful of what is happening around oneself to understand how information, events, and one's own actions will impact one's goals and objectives; another key component of a HRO.
- We approached checklist use first by facilitating the centers' response to the mandate by the California Department of Health Care Services to comply with the Centers for Disease Control and Prevention's (CDC's) NHSN ''central line insertion process.''[45] We encouraged units to find other opportunities for the use of checklists; such as for line set-up, line entry, and medication administration.
- We approached ''error interception,'' often referred to by its name: ''stop the line'' as originally conceived of and termed by Toyota Motor Inc.,[46] as a key characteristic of a safety culture.[47] We built on the national effort to improve the safety of line insertion, a simple instance of events that rely upon the staff's comfort level in stopping a process until their safety concerns are addressed,[48] as a first step to embracing this habit for safe care.
- We addressed the need for situational awareness by encouraging the use of huddles to identify critical issues in each unit before each shift.
- We promoted the habit of reflection and feedback about misses and near misses by transforming the ''root cause analysis'' process for investigating possible CLABSI events from a retrospective into a concurrent process. The concurrent process changes the analytic emphasis from intense chart review and cause documentation to bedside interview and observation. As it evolved, the positive blood culture evaluation process pulled together both the experiences of our members as well as published observations on those factors relevant to understanding the circumstances surrounding each CLABSI.
- Last, we developed auditing tools for monitoring critical processes. While this has widely been recognized and implemented for hand hygiene, we developed audit tools for line set-up and line entry with the intent that units would periodically monitor these critical performance items in the same way that they monitory hand hygiene. (Additional forms and materials at the project's Web site: http://www.dhcs.ca.gov/ProvGovPart/initiatives/nqi/Pages/default.aspx. Accessed on January 14, 2009.) (**Figs. 7** and **8**).

Accounting for the cultural dimension of neonatal ICU care

We developed a ''metabundle'' of recommended interventions (**Table 3**) to distinguish the adaptive or social context surrounding the implementation of the bundles'

Table 2
CCS-CCHA neonatal CLABSI prevention bundle (2008): technical aspects: http://www.dhcs.ca.gov/ProvGovPart/initiatives/nqi/Pages/default.aspx

Performance Expectations	Considerations
Insertion	
1. Maximum sterile barrier precautions used	- Cover entire infant with sterile drapes or as much as affords safe observation. - Recommend staff wear face mask when within 3 ft of sterile field
2. Skin disinfected with Chlorhexidine (CHG) or povidone iodine (PI)	- Apply over 30 seconds & allow to dry (exception aqueous CHG)
3. Dedicated team for placement & maintenance	- Insertion training course, including sterile technique, hand hygiene, use of maximum sterile barrier precautions, proper skin disinfection - Educational competencies for all aspects of care
4. All supplies required for the procedure should be available at the bedside before catheter insertion	—
5. Hand hygiene standards met	—
6. Insertion checklist used	- Standardize critical elements of line insertion - Ensure staff observers are skilled in monitoring elements of sterile technique.
7. Staff empowered to stop non-emergent procedure if sterile technique not followed	—
Maintenance	**Considerations**
Assessment & Site Care	
1. Daily assessment and documentation of catheter need included as part of multidisciplinary rounds and review of daily goals	When catheter used primarily for nutritional purposes: - Consider removal when infant reaches $\geq$120 mL/kg/day enteral nutrition - Consider discontinuing lipids when infant reaches >2.5 g/kg/day of enteral fat intake

2. Review dressing integrity and site cleanliness daily	Change PRN using sterile technique and CHG or PI for skin antisepsis
Tubing, injection ports, catheter entry	
1. Use "closed" systems for infusion, blood draws & medication administration	- May use manufactured or improvised closed system. If stopcocks are used, port(s) are capped with swabable needleless connector(s). - Define consistent practice to be used when accessing catheters
2. Assemble and connect infusion tubing using aseptic or sterile technique. Configure tubing consistently for each type of VAD.	- **Sterile** technique ideally includes sterile barrier for tubing assembly and wearing of face mask, hat, sterile gloves & 2 staff members performing connection to central catheter - **Aseptic** technique includes clean barrier for tubing assembly & wearing of clean gloves
3. Scrub needleless connector using friction with either alcohol or CHG/alcohol swab for at least 15 sec. before entry. Allow surface to dry before entry.	—
4. Clean gloves for all VAD entries & hand hygiene used before & after glove use	Standard precautions
5. Use pre-filled, flush containing syringes wherever feasible	- Higher risk of contamination when flush withdrawn from another container by a nurse
6. Staff empowered to stop non-emergent procedure if sterile technique not followed	—

Positive Blood Culture Review
Version 7: 6/30/08

Name: _____ MR #: _____ DOB:__/__/__ Birth WT:_____ GA_____wk EGA at dx:____wk

Final Dx: [ ] CABSI [ ] BSI [ ] NEC [ ] VAP [ ] Contaminant [ ]

Other:_____

Clinical Findings at Time Blood Culture Drawn and Blood Culture Collection Data	
Risk Factors: [ ] Immunocompromised [ ] Compromised skin integrity [ ] Open body cavity [ ] Ostomy present [ ] Surgical site infection receiving Rx [ ] Other risk factors: (state)	[ ] Blood transfusion in last 72 hours [ ] NCPAP/Nasal cannula [ ] Feeding tube Enteral nutrition volume/parenteral volume ratio: _____ (approximate) [ ] Vascular Catheter [ ] Major surgery within past week [ ] or any other time Specify type of surgery:
Catheter Information: **Only relevant if line(s) present within the 48 hr prior to first blood culture**	
[ ] None [ ] PIV ____# days (if multiple site, note only longest) Estimate # IV start attempts in last 72 hrs:_____ [ ] UAC_____# days [ ] UVC_____# days [ ] PICC_____# days Site: _____ [ ] Other CENTRAL line _____# days Site: _____ _____Estimate total # times all lines accessed during the **last 72 hours** (including for meds/blood draws/ tubing changes, etc) Last date dressing changed: ___/___/___	[Y/N] Antibacterial patch in use [Y/N] Abnormal CL site appearance on day culture drawn [Y/N] Line-related phlebitis [Y/N] Compromised dressing [Y/N] Vomiting onto line dressing [Y/N] Stool/Urine onto line dressing [Y/N] Line repaired/exchanged in past 48 hours [Y/N] Line leaking events [ ] Care by temporary staff [ ] Care by non-NICU staff [ ] Staffing difficulties [ ] Improper line set-up [ ] Tubing/infusate NOT changed appropriately (method/time) [ ] Any other unusual event: (specify)
Infusates in Past 72 hours : [ ] TPN [ ] Lipids [ ] Blood products [ ] Steroids (3 x physiologic doses)	
Comments and Lessons Learned:	[ ] BSI (Not a CLA-BSI) after further review (e.g., meets another CDC definition and there is another clear source identified), e.g. [ ] NEC [ ] VAP [ ] Other:_____
	[ ] CLA-BSI (have data to determine that BSI fulfills CLA-BSI criteria, i.e. BSI very likely related to CL) *See page 2 for definition detail*
	[ ] Contaminant
	Adjudication Process:
	[ ] BSI Event was clearly able to be attributed/categorized into a CDC definition [ ] BSI Event required significant inferences/judgment to be attributed/categorized
	Action Plan: (Please relate to Fishbones, as applicable):

Fig. 6. Positive blood culture evaluation process (only summary page shown; remaining 3 pages available at: http://www.dhcs.ca.gov/ProvGovPart/initiatives/nqi/Pages/default.aspx. Accessed January 14, 2009).

technical features. It identifies diagnostic criteria; addresses leadership and other structural aspects of the bundle's implementation, specifies use of HRO processes including a preoccupation with failure and creating a culture of safety, the periodic and routine monitoring of processes, reflection on adverse events and feeding back of results; and encourages interdepartmental exchanges to ensure bundle compliance while the neonate is receiving care in other departments of the hospital.

This collaborative discovered the importance of context when implementing quality improvement practices. "*When you have seen one NICU…you have seen one NICU*" became a maxim of our collaborative. Although all 22 sites met the same definition of "regional NICU," there were very significant differences in staffing, the role of house staff and fellows, the mix of surgical versus VLBW neonates, other initiatives on-going in the NICU which might distract attention and resources from the CLABSI project, and

IV Tubing Change Procedure Observations

Date: _____ Shift: _____ Observer: _____

Observation #	1	2	3	4	5	6	7	8	9	10
Perform hand hygiene adequately	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N
Gather supplies	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N
Use disinfectant wipe to clean surface of counter	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N
Clean or sterile barrier used for tubing assembly	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N
Place new tubing & supplies on barrier without contamination	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N
Perform hand hygiene adequately	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N
Wear sterile/clean gloves during tubing assembly	Y N NA	Y N NA	Y N NA	Y N NA	Y N NA	Y N NA	Y N NA	Y N NA	Y N NA	Y N NA
Wear items required for procedure, per facility P&P H=hair covering F=face mask G=gown	H F C NA	H F C NA	H F C NA	H F C NA	H F C NA	H F C NA	H F C NA	H F C NA	H F C NA	H F C NA
Assemble tubing using all required components & without contamination	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N
Connect IV solution to tubing & prime without contamination	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N
Place tubing in bed without contaminating end	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N
Place 4X4 under CVC connection site	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N
Scrub connections 10-15 sec with alcohol/CHG before disconnecting	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N
Perform hand hygiene & don clean/sterile gloves	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N
Connect tubing to VAD without contamination	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N
Perform hand hygiene after glove removal	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N
Label tubing with date & time	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N
Provide feedback	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N	Y N
Comments										

Fig. 7. Intravenous tubing change procedure audit tool.

Catheter Entry Observations

Date: Shift: Observer:

Observation #	1	2	3	4	5
Reason for entry: B=blood draw T=tubing change M=med administration	B T M	B T M	B T M	B T M	B T M
Type of catheter P=PIV, UA=UAC; UV=UVC; PI=PICC O=Other	P UA UV PI O	P UA UV PI O	P UA UV PI O	P UA UV PI O	P UA UV PI O
Perform hand hygiene	Y	Y	Y	Y	Y
	N	N	N	N	N
Apply clean gloves	Y	Y	Y	Y	Y
	N	N	N	N	N
Place clean/sterile barrier under port	Y N NA	Y N NA	Y N NA	Y N NA	Y N NA
Scrub port for 10-15 sec using friction	Y	Y	Y	Y	Y
	N	N	N	N	N
Port scrubbed using: A = Alcohol C = Chlorhexidine	A C	A C	A C	A C	A C
Allow anti-microbial to dry	Y	Y	Y	Y	Y
	N	N	N	N	N
Enter port without contamination	Y	Y	Y	Y	Y
	N	N	N	N	N
Flush obtained from vial without contamination	Y N NA	Y N NA	Y N NA	Y N NA	Y N NA
Single use prefilled flush syringe used only 1 time	Y N NA	Y N NA	Y N NA	Y N NA	Y N NA
All blood/residue flushed clear from injection port	Y N NA	Y N NA	Y N NA	Y N NA	Y N NA
Wipe blood from port surface following blood draw	Y N NA	Y N NA	Y N NA	Y N NA	Y N NA
Remove gloves	Y	Y	Y	Y	Y
	N	N	N	N	N
Perform hand hygiene	Y	Y	Y	Y	Y
	N	N	N	N	N
Provide feedback	Y	Y	Y	Y	Y
	N	N	N	N	N
Comments					

Fig. 8. Catheter entry audit tool.

the staff and technical resources available to undertake such an effort. The project team rigorously maintained focus on the intent of the project, to reduce CLABSIs, but permitted significant freedom in how individual sites tried to achieve that goal. Also emphasized was the critical role played by the unit's culture of safety and how that is manifested in the interpersonal, interdisciplinary relationships within the unit.

Implementation issues included: whether, and how, to implement feedback to the staff about days without an infection; whether or not to allow parents to become aware of these data; how to involve the parents in the hand hygiene campaigns, even to the extent of empowering them to "stop the line" if they had concerns; and how to approach nonconforming colleagues.

We addressed communication problems and the possibility of infections attributable to line manipulation outside of the NICU by encouraging dialog with non-NICU departments such as Anesthesiology and Radiology, who also need access to these lines but are variably educated and trained to the same standards for their set up and entry. We proactively sought improved communication with Infection Control Departments, whose focus is on ensuring each report's completeness and accuracy and whose routine reporting lag can delay the timely assessment of the clinical situation surrounding an event and its feedback to the NICU. We encouraged NICU leaders to interact with their Hospital's leadership about their efforts and results to reduce CLABSIs, as these leaders were now being addressed externally by the Joint Commission, the Centers for Medicare and Medicaid Services and state initiatives to publicly release adverse event rates.[49,50] We encouraged the NICU teams to address staff communication challenges within their units by adopting techniques such as "huddles" at shift change; immediate bedside consultation with and by the staff as soon as any positive blood culture report was received; and "rounding for outcomes."

Summary, Part 2

In our experience, improvement collaboratives must have two closely linked goals. The first is to improve current performance such as reducing the rate of CLABSIs by 50% over 6 months. The second goal must be to sustain those improvements over time even in the face of competing priorities and the never-ending distractions of everyday clinical care. To achieve the first goal, the California collaborative emphasized the technical aspects of care such as proper line set up and medication administration supported by appropriate bundles, checklists, and audits to encourage clinicians to do the right thing. To help sustain the gains, which is often much harder than making the initial improvements, we focused on the adaptive changes[51] the social and cultural aspects of interdisciplinary care within the NICU including the need for open communication with staff (and families); open questioning by any staff member of the usual way of doing things; continuous learning; appropriate modeling of behavior by leaders; and a focus on failures and ways to prevent, identify and mitigate them.

Our experience highlights the notion that productive improvement collaboratives entail fundamental and transformative change in the way providers deliver and evaluate their care. Far from mere platitude, the point is that "every system is perfectly designed to achieve the results that it gets" (attributed to Deming, by Pauker and colleagues[52]) so to get better results requires changing the system. California and New York neonatologists have learned that achieving and sustaining such change is a monumental undertaking; demanding affirmation that the current reality is unacceptable, a clear and shared vision of the desired future, and dogged determination – over many years.

Table 3
CCS-CCHA neonatal central line associated bloodstream infection prevention bundle (2008): adaptive aspects

Administrative Leadership	Considerations
1. Demonstrable administrative involvement in and support for achieving Zero Health care–Associated Infections	—
2. Engage Staff with feedback: Posting days since last CABSI Posting CABSI rates	- Annotate CABSI rates with descriptions and dates of practice changes - Celebrations of successes
3. Perform investigation and analysis of each CABSI	- Begin process ASAP & within 24 hours of CABSI notification. Review opportunities for system improvements after each event.
4. Surveillance activities of critical processes related to sustaining the gains: a. Hand Hygiene b. Adherence to unit catheter management and entry standards c. Monitor patient processes off unit for bundle compliance d. Unit personnel support for the "Stop the Line" safety culture	a. Capture 50 HH observations/month/activity using consistent observers b. As above initially, then smaller volume less frequently. c. Prospectively establish and maintain bundle compliance with off unit service departments, eg, operating rooms (Anesthesiology and Pediatric Surgery), radiology suite (Radiology). d. Empower staff to stop intervention at any time when technique is being breached.
5. Competent trained personnel to perform specialized maintenance activities	- Consider specialized team for dressing changes, catheter repair, catheter clearance of blockage
CABSI Diagnosis and Classification	
1. Two blood cultures drawn from separate sites, following skin disinfection with PI or CHG, within 48 hours of each other.	- One culture may be from a central line site if a second peripheral site is not feasible, taking into account circumstances such as vessel accessibility, pain and the infant's clinical status. - The recommended neonatal culture volume is $\geq$1 mL

2. The diagnosis of a laboratory confirmed (LC) catheter-associated BSI (CABSI) can only be made in the absence of another clinically appreciated infectious focus, the presence of one or more positive blood cultures, and one of the following three criteria being met:

Criteria 1: at least one blood culture growing a recognized pathogen (see Considerations); or

Criteria 2: at least two blood cultures growing a recognized contaminant (see Considerations) and the presence of one (or more) clinical signs of generalized infection (either Fever >38°C (see Considerations) or Hypotension; or

Criteria 3: Age <1y AND one of the following:

Fever (see Considerations), Hypothermia (<37°C rectal), apnea, or bradycardia.

See: http://www.cdc.gov/ncidod/dhqp/pdf/nhsn/NHSN_Manual_PatientSafetyProtocol_CURRENT.pdf

- Recognized pathogens are those not named as common skin contaminants.
- Common skin contaminants: diphtheroids, Bacillus species, Propionibacterium species, coagulase-negative staphylococci including *S epidermidis*, viridans group streptococci, Aerococcus or Micrococci
- Fever: per the CDC's NHSN, the neonatal equivalents of >38°C rectal are: (38°C rectal/tympanic/temporal art 37°C oral = 36°C axillary)
- Hypotension is not defined further.
- Hypothermia: per the CDC's NHSN, the neonatal equivalents of <37°C are: (37°C rectal/tympanic/temporal artery = 36°C oral = 35°C axillary)

However this collaborative does not believe the temperature equivalences specified by NHSN realistically reflect their neonatal populations' temperature data.

Instead the collaborative recommends that **axillary temperatures should be considered as a screening method; axillary temperatures** < 36.0°C (<96.8°**F) should be tentatively labeled as "hypothermia" and axillary temperatures › 38.0°C (**>100.4°**F) should be tentatively labeled as fever.** *Because of the variability in axillary temperature readings, the presence of an elevated or hypothermic temperature will only be termed* ***confirmed*** if there have been ***at least two consecutive abnormal axillary measurements or one abnormal axillary and one abnormal rectal (or other core) measurement.***

Available at: http://www.dhcs.ca.gov/ProvGovPart/initiatives/nqi/Pages/default.aspx.

REFERENCES

1. Kruger J, Dunning D. Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. J Pers Soc Psychol 1999;77(6):1121–34.
2. Wikipedia. Lake Wobegon effect. Available at: http://en.wikipedia.org/wiki/Lake_Wobegon_effect; August 9, 2006. Accessed August 10, 2008.
3. Davis DA, Mazmanian PE, Fordis M, et al. Accuracy of physician self-assessment compared with observed measures of competence. JAMA 2006;296(9):1094–102.
4. Chassin MR, Galvin RW. National Roundtable on Health Care Quality. The urgent need to improve health care quality. JAMA 1998;280(11):1000–5.
5. Vermont Oxford Network. The Vermont Oxford Network. Available at: http://www.vtoxford.org/. 2008. Accessed July 24, 2008.
6. Hennekens CH, Buring JE. Epidemiology in medicine. Boston: Little, Brown and Company; 1987.
7. Schulman J, Spiegelhalter DJ, Parry G. How to interpret your dot: decoding the message of clinical performance indicators. [State of the Art Article]. J Perinatol 2008;28:588–96.
8. Ash AS, Schwartz M. Evaluating the performance of risk-adjustment methods: dichotomous outcomes. In: Iezzoni LI, editor. Risk adjustment for measuring healthcare outcomes. Chicago: Health Administration Press; 1997. p. 427–70.
9. Schulman J. Managing your patients' data in the neonatal and pediatric ICU: an introduction to databases and statistical analysis. Oxford, UK: Blackwell; 2006.
10. van Belle G. Statistical rules of thumb. In: Balding DJ, Bloomfield P, Cressie NA, editors. Wiley series in probability and statistics. New York: Wiley-Interscience; 2002. p. 221.
11. Schulman J. Transforming risk factors into an unbiased mortality model. J Perinatol 2008;28:243–6.
12. Schulman J, Stricof RL, Stevens TP, et al. Development of a Statewide Collaborative to decrease NICU central line associated bloodstream infections. J Perinatol, in press.
13. Bureau of Women's Health. N.Y.S.D.O.H. NYS Hospitals and perinatal designation level. Available at: http://www.health.state.ny.us/community/pregnancy/health_care/perinatal/hospital_designations.htm. September, 2005. Accessed June 30, 2008.
14. Division of Healthcare Quality Promotion and National Center for Infectious Diseases. The National Healthcare Safety Network (NHSN) user manual. Department of Health and Human Services, Centers for Disease Control and Prevention; 2006.
15. Centers for Disease Control and Prevention. Guidelines for the prevention of intravascular catheter-related infections. MMWR Recomm Rep 2002;51(RR-10):1–29.
16. Berenholtz SM, Pronovost PJ, Lipsett PA, et al. Eliminating catheter-related bloodstream infections in the intensive care unit. Crit Care Med 2004;32:2014–20.
17. Provonost P, Needham D, Bernholtz S, et al. An intervention to decrease catheter-related bloodstream infections in the ICU. N Engl J Med 2006;355(26):2725–32.
18. Centers for Disease Control and Prevention. Guideline for hand hygiene in healthcare settings: recommendations of the Healthcare Infection Control Practices Advisory Committee and the HICPAC/SHEA/APIC/IDSA Hand Hygiene Task Force. MMWR Recomm Rep 2002;51(RR-16):1–45.

19. Boyce JM, Pittet D. Guideline for hand hygiene in health-care settings: recommendations of the Healthcare Infection Control Practices Advisory Committee and the HICPAC/SHEA/APIC/IDSA Hand Hygiene Task Force. Infect Control Hosp Epidemiol 2002;23(Suppl 12):S3–40.
20. Raad II, Hohn DC, Gilbreath BJ, et al. Prevention of central venous catheter-related infections by using maximal sterile barrier precautions during insertion. Infect Control Hosp Epidemiol 1994;15(4 Pt 1):231–8.
21. Mimoz O, Villeminey S, Ragot S, et al. Chlorhexidine-based antiseptic solution vs alcohol-based povidone-iodine for central venous catheter care. Arch Intern Med 2007;167(19):2066–72.
22. Chaiyakunapruk N, Veenstra VL, Lipsky BA, et al. Chlorhexidine compared with povidone-iodine solution for vascular catheter–site care: a meta-analysis. Ann Intern Med 2002;136:792–801.
23. McGee DC, Gould MK. Preventing complications of central venous catheterization. N Engl J Med 2003;348:1123–33.
24. Maki DG. A prospective, randomized trial of gauze and two polyurethane dressings for site care of pulmonary artery catheters: implications for catheter management. Crit Care Med 1994;22:1729–37.
25. Gillies D, O'Riordan L, Carr D, et al. Gauze and tape and transparent polyurethane dressings for central venous catheters. Cochrane Database Syst Rev 2003;(3):1–19.
26. Crnich CJ, Maki DG. The promise of novel technology for the prevention of intravascular device–related bloodstream infection. I. Pathogenesis and short-term devices. Clin Infect Dis 2002;34:1232–42.
27. Salzman MB, Isenberg HD, Shapiro JF, et al. A prospective study of the catheter hub as the portal of entry for microorganisms causing catheter-related sepsis in neonates. J Infect Dis 1993;167(2):487–90.
28. Aly H, Herson V, Duncan A, et al. Is bloodstream infection preventable among premature infants? A tale of two cities. Pediatrics 2005;115:1513–8.
29. Garland JS, Alex CP, Sevallius JM, et al. Cohort study of the pathogenesis and molecular epidemiology of catheter-related bloodstream infection in neonates with peripherally inserted central venous catheters. Infect Control Hosp Epidemiol 2008;29:243–9.
30. Association for Professionals in Infection Control and Epidemiology, I., Healthcare-Associated Reporting Laws and Regulations. 2008.
31. McGiffert L. Summary of state activity on hospital-acquired infections. Austin, Texas: Consumers Union; 2007. Available at: http://www.consumersunion.org/campaigns/CU%20Summ%20HAI%20state%20rpting%20laws%20as%20of%201-08.pdf.
32. Department of Health. N.Y.S., New York state hospital-acquired infection reporting system: pilot year–2007 2008. Available at: http://www.nyhealth.gov/nysdoh/hospital/reports/hospital_acquired_infections/2007/docs/hospital-acquired_infection-full_report.pdf. Albany (NY).
33. Pronovost PJ, Goeschel CA, Wachter RM. The wisdom and justice of not paying for "preventable complications". JAMA 2008;299(18):2197–9.
34. Costello JM, Morrow DF, Graham DA, et al. Systematic intervention to reduce central line-associated bloodstream infection rates in a pediatric cardiac intensive care unit. Pediatrics 2008;121(5):915–23.
35. Pawson R, Tilley N. Realistic evaluation. London: Sage Publications, Ltd; 1997.
36. Horn SD. Performance measures and clinical outcomes. JAMA 2006;296(22): 2731–2.
37. Berwick DM. The Science of improvement. JAMA 2008;299(10):1182–4.

38. Schulman J. Evaluating the processes of neonatal intensive care. London: BMJ Books; 2004.
39. Benneyan J, Lloyd R, Plsek P. Statistical process control as a tool for research and healthcare improvement. Qual Saf Health Care 2003;12:458–64.
40. Edwards JR, Peterson KD, Andrus ML, et al. National Healthcare Safety Network (NHSN) report, data summary for 2006, issued June 2007. Am J Infect Control 2007;35:290–301.
41. Division of Healthcare Quality Promotion. The National Healthcare Safety Network (NHSN) manual: patient safety component protocol. Atlanta (GA): National Center for Infectious Diseases; 2008.
42. Bizzarro M, Raskind C, Baltimore RS, et al. Seventy-five years of neonatal sepsis at Yale: 1928–2003. Pediatrics 2005;116:595–602.
43. Luria JW, Schoettker PJ. Reliability science and patient safety. Pediatr Clin North Am 2006;53:1121–33.
44. Studer Q. Hardwiring excellence. Gulf Breeze (FL): Fire Starter Publishing; 2003.
45. National Healthcare Safety Network (NHSN). Central Line Insertion Practices (CLIP) adherence monitoring. Available at: http://www.cdc.gov/ncidod/dhqp/nhsn_CLIP_AdherenceMonitoring.html. Accessed December 22, 2008.
46. Glauser E. The Toyota phenomenon, 2005.
47. ECRI Institute (2005) Health Care Risk Control-Culture of Safety. Risk Management Reporter Volume.
48. Pronovost P, Berenholtz SM, Goeschel CA, et al. Creating high reliability in health care organizations. Health Serv Res 2006;41:1599–617.
49. HAI prevention emphasized in 2009 National Patient Safety Goals. Healthcare Benchmarks Qual Improv 2008;85–8.
50. Jarvis W. The Lowbury lecture. The United States approach to strategies in the battle against healthcare-associated infections, 2006: transitioning from benchmarking to zero tolerance and clinician accountability. J Hosp Infect 2007;65(Suppl 2):3–9.
51. Goeschel C, Pronovost P. Harnessing the Potential of Health Care Collaboratives: lessons from the keystone ICU project, in Advances in patient safety: new directions and alternative approaches. Volumes 1–4, AHRQ Publication Nos. 08-0034 (1–4). July 2008. Rockville (MD): Agency for Healthcare Research and Quality; 2005.
52. Pauker S, Zane E, Salem D. Creating a safer health care system: finding the constraint. JAMA 2005;294(22):2906–8.