SECTION 3: CASE STUDIES

The Vermont Oxford Network: Evidence-Based Quality Improvement for Neonatology

Jeffrey D. Horbar, MD

ABSTRACT. The Vermont Oxford Network is a voluntary collaborative group of health professionals committed to improving the effectiveness and efficiency of medical care for newborn infants and their families through a coordinated program of research, education, and quality-improvement projects. In support of these activities, the Network maintains a clinical database of information about very low birth weight infants that now has more than 300 participating neonatal intensive care units (NICUs). We anticipate that these NICUs will submit data for 25 000 infants with birth weights of 401 to 1500 g born in 1998.

The research program of the Network includes outcomes research and randomized clinical trials. The goal of Network outcomes research is to identify and explain the variations in clinical practice and patient outcomes that are apparent among NICUs. Network trials are designed to answer practical questions of importance to practitioners and families using pragmatic designs that can be integrated into the daily practice of neonatology. Quality improvement is a major focus of the Network. Members receive confidential quarterly and annual reports based on the Network database that document their performance and compare practices and outcomes at their unit with those at other units within the Network. These reports are intended to assist the members in identifying opportunities for improvement and to help them monitor the success of their improvement efforts.

Although information is necessary for improvement to occur, it is not sufficient to foster lasting improvement by itself. Information must be translated into action. The Network is sponsoring an ongoing program of quality initiatives designed to provide members with the knowledge, skills, tools, and resources needed to foster action for improvement.

The Network’s first formal quality-improvement project, the NIC/Q Project, brought together 10 NICUs to apply the methods of collaborative improvement and benchmarking to neonatal intensive care. Building on the lessons learned in that initial project, the Network now is conducting the Vermont Oxford Network Evidence-Based Quality Improvement Collaborative for Neonatology, known as NIC/Q 2000. This 2-year collaborative will assist multidisciplinary teams from the 34 participating NICUs to develop four key habits for improvement: the habit for change, the habit for practice as a process, the habit for collaborative learning, and the habit for evidence-based practice. During the collaborative, participants will contribute to a knowledge bank of clinical, organizational, and operational change ideas for improving neonatal care.

The coordinated program of research, education, and quality improvement described in this article is only possible because of the voluntary efforts of the members. The Network will continue to support these efforts by developing and providing improved tools and resources for the practice of evidence-based neonatology. Pediatrics 1999;103:350–359; neonatology, very low birth weight, database, network, quality improvement, evidence-based medicine, randomization, trials, outcomes, mortality, length of stay.

ABBREVIATIONS. NICU, neonatal intensive care unit; NICHD, National Institute of Child Health and Development; LOS, length of stay; OR, odds ratio; CI, confidence interval; SMR, standardized mortality ratio; NIC/Q, Neonatal Intensive Care Collaborative Quality (Project); PDSA, Plan-Do-Study-Act; CAT, critically appraised topic.

The Vermont Oxford Network was established in 1989 with the goal of improving the effectiveness and efficiency of medical care for newborn infants and their families through a coordinated program of research, education, and quality-improvement projects. This article will describe the Network and its activities, with a particular focus on the Vermont Oxford Network Evidence-based Quality Improvement Collaborative for Neonatology.

The Vermont Oxford Network is a nonprofit corporation supported by membership fees, grants, and contracts. The basic philosophy of the Network is to integrate research into daily practice by designing simple, pragmatic studies that are compatible with the demands of busy health professionals and relevant to the questions that arise in daily practice.1 Our philosophy is based on the concepts developed by Ian Chalmers, Adrian Grant, and co-workers at the National Perinatal Epidemiology Unit in Oxford, England. The Network has adopted these ideas and, in so doing, gained the cooperation and allegiance of neonatologists, neonatal nurses, and other professionals who volunteer their time and effort to participate in the Network. They are motivated both by a desire to contribute to new knowledge regarding newborn medicine and to have access to comparative
Network data that allows them to evaluate their own performance and that of their NICU.

VERMONT OXFORD NETWORK DATABASE

The Vermont Oxford Network maintains a database including information for all infants with birth weights of 401 to 1500 g born at member institutions or admitted to them within 28 days of birth (before 1996, only infants weighing 501 to 1500 g were included). The database provides information for use in outcomes and epidemiologic research, serves as the core data source for Network trials, and is used to provide comprehensive individualized reports to participating hospitals that serve as the foundation for local quality-improvement projects and peer review. The database is a unique source of information regarding clinical practices and patient outcomes for high-risk infants and is successfully supporting the mission of the Network. It has generated numerous scientific presentations, reports, and publications.1,2,4-16

The database has grown dramatically during the past 9 years (Fig 1). In 1990, the first full year of database operations, 36 hospitals submitted data for ~3000 very low birth weight infants. The Network now has more than 300 member NICUs. In 1998, we anticipate collecting data for close to 25 000 very low birth weight infants. This will include >50% of all very low birth weight infants born in the United States in that year. The Network currently is performing a pilot project designed to evaluate the feasibility of data collection for all NICU admissions, regardless of birth weight. We anticipate introducing this expanded database to our members over the next 2 years.

RANDOMIZED CLINICAL TRIALS

An important component of the Vermont Oxford Network’s mission is to generate new knowledge and evidence for practice by performing randomized controlled trials. The Network clinical trials program is directed by Roger F. Soll, MD. Network trials are supervised by steering committees composed of Network members and are monitored by appointed data safety committees. Investigators from member institutions volunteer their time and effort to perform the trials.

The first Network multicenter trial compared two surfactants for the treatment of neonatal respiratory distress syndrome.2,17 This trial enrolled >1200 very low birth weight infants at 38 Network sites. It achieved results comparable to those of a smaller trial of similar design performed by the National Institute of Child Health and Development (NICHD) Neonatal Research Network, a grant-supported network of academic research centers.18 The surfactant trial was an excellent choice for the first Vermont Oxford Network trial, because it did not involve an experimental therapy. Both of the surfactants studied were available commercially at the time the trial was performed, allowing the investigators to focus on the mechanics of trial participation including identifying eligible subjects, obtaining informed consent, randomizing assignment of treatments, and collecting trial data. The successful completion of the surfactant trial demonstrated the willingness and capability of volunteer Network investigators to perform large, high-quality randomized trials at a low cost.

The Network currently is completing a randomized trial to determine whether corticosteroid treatment within 12 hours of birth decreases mortality and chronic lung injury for infants weighing <1000 g at birth. The trial, with a planned sample size of 800 infants at more than 40 centers, is a major milestone for the Network, because it requires the pharmacies at participating sites to prepare coded vials of medication and to dispense them according to a locally maintained randomization lists. The pharmacists at these institutions are volunteering their time and effort and receive no financial support from the Network. The participation of pharmacists at each site will allow the Network to perform a wide range of masked drug trials. It provides additional evidence of the commitment and ability of member hospitals to integrate research into the daily practice of neonatal intensive care.

The Network now is preparing to begin a trial of skin care for the prevention of nosocomial infection in extremely low birth weight infants (William H. Edwards, MD, Principal Investigator; Jeannette Connor, MS, MN, ARNP, Co-investigator, Dartmouth Hitchcock Medical Center). The primary goal of the study is to evaluate the impact of Aquaphor Original emollient ointment treatment during the first 2 weeks of life on the incidence of mortality and/or nosocomial bacterial sepsis for infants with birth weights of 501 to 1000 g. The premise of this trial is that skin breakdown and the resulting loss of integrity of the skin barrier lead to an increased risk of nosocomial infection. Data from one small trial suggest that this may be the case.19 The Network trial will enroll 800 infants and have sufficient power to detect a 10% absolute risk reduction (from 40% to 30%) in the combined rate of nosocomial infection or death in the target population.

Network trials are part of a coordinated strategy of research and quality improvement. Both the early corticosteroid trial and the skin care trial have direct

Fig 1. The number of NICUs participating in the Vermont Oxford Network Database each year (left) and the number of infants born and enrolled in the Database each year (right) from 1990 to 1998. The data for 1998 are projections.
OUTCOMES RESEARCH

Dramatic variations in clinical practice and patient outcomes have been observed in a wide variety of medical settings. Practices and outcomes vary among different physicians, geographic regions, and health care institutions. The goal of outcomes research is to identify and explain these variations. The Vermont Oxford Network Database demonstrates the presence of substantial variation among individual NICUs with respect to mortality, morbidity, and length of stay (LOS) for very low birth weight infants. Analyzing this variation, both over time and among different NICUs, has served as a focus for outcomes research in the Network.

The database has been used to evaluate the effect of antenatal steroid treatment on patient outcomes, to identify trends in morbidity, mortality, and clinical practices during the 1990s, and to determine patient and hospital characteristics associated with the mortality risk. Studies are in progress concerning delivery room resuscitation practices, the effects of intrauterine growth retardation on morbidity and mortality for premature infants, and the association of different levels of NICU care with outcomes.

Network studies of antenatal corticosteroid treatment provide a good example of the how the database can be used for outcomes research and how the results of this research can be applied. In February 1994, the National Institutes of Health convened the Consensus Development Conference on the Effect of Corticosteroids for Fetal Maturation on Perinatal Outcomes. The Consensus Development Panel, in preparation for this conference, invited several organizations with neonatal databases to provide data for presentation. The organizations (Burroughs Wellcome, the NICHD Neonatal Research Network, Ross Laboratories, and the Vermont Oxford Network) worked together to develop a common analytic plan and format for presentation.

During 1991 and 1992, 73 centers participated in the Network database. These centers submitted data for 8908 infants with birth weights of 501 to 1500 g. After excluding infants with missing data (n = 32) and those who died in the delivery room (n = 127), the final sample for analysis included 8749 infants. Seventy-four percent of these infants were not exposed to any antenatal corticosteroid treatment, 7.5% were exposed to partial treatment (delivery <24 hours or >1 week after the last dose of maternal steroid treatment), and 18.5% were exposed to a complete course of treatment (delivery between 24 hours and 1 week of the last dose of maternal steroid treatment). A logistic regression analysis showed that any corticosteroid treatment (partial or complete) was associated with reductions in the risks of death within 28 days of birth (odds ratio [OR] = 0.55; 95% confidence interval [CI]: 0.45 to 0.64), respiratory distress syndrome (OR = 0.62; 95% CI: 0.56 to 0.70), intraventricular hemorrhage (OR = 0.71; 95% CI: 0.59 to 0.76), and severe intraventricular hemor-

rhage (OR = 0.60; 95% CI: 0.45 to 0.69). Antenatal corticosteroid exposure was associated with an increased risk of necrotizing enterocolitis (OR = 1.32; 95% CI: 1.08 to 1.60). There was no statistically significant association with the risk of sepsis (OR = 1.10; 95% CI: 0.93 to 1.20). The results of this analysis were consistent with those of the other organizations presenting data to the Consensus Development Panel.

Broad recommendations regarding the use of antenatal corticosteroids for women at risk for preterm delivery are justified by the evidence from randomized trials alone. However, because neonatal intensive care practices have changed dramatically since the trials were conducted and because additional further trials are unlikely to be performed, the observational data provided by the Vermont Oxford Network and the other participating organizations added supplementary evidence of value for the panel to use in its deliberations.

Additional Network research has shown that antenatal corticosteroid use has increased steadily since 1990. The overall percentage of infants weighing 501 to 1500 g in the Network Database exposed to any antenatal corticosteroid therapy increased from 19% in 1990 to 34% in 1993. During that period, increasing year of birth, prenatal care, inborn location of birth, and multiple birth were associated with a higher likelihood of antenatal corticosteroid exposure, whereas black race and small size for gestational age were associated with a lower likelihood of exposure.

Antenatal corticosteroid treatment has continued to increase in subsequent years. In 1996, the median rate of antenatal corticosteroid treatment for infants weighing 501 to 1500 g at 191 Network hospitals was 66%. However, there was substantial variation in these rates among hospitals. The one quarter of the hospitals with the highest rates had rates >76%, whereas the one quarter with the lowest rates had rates of ≤56%. It is likely that hospitals in the upper quartile for antenatal corticosteroid treatment rates are now providing treatment to most eligible women, whereas those in the lowest quartile still could increase their treatment rates.

Preliminary analyses of the database presented in abstract form indicate that mortality rates for infants weighing 501 to 1500 g have decreased steadily since 1991 in association with the increased use of antenatal corticosteroids. There have been a number of other changes in obstetric and neonatal practice, as well as changes in the expectations of physicians and parents during that period that could account for this observation. However, the evidence suggests that the increase in antenatal corticosteroid treatment for women at risk for preterm delivery is one important contributing factor. If that is the case, then it may be that as treatment rates increase to the point that all or most eligible women are receiving antenatal corticosteroid treatment, the observed decline in mortality will slow or level off.

It is important to recognize that in addition to providing new information to the scientific community, as in the corticosteroid example, the results of Network research allow us to design and create more

352 SUPPLEMENT
useful reports for the members. These reports are intended to help the members identify areas in which they have opportunities for improvement. To do this, the reports must compare “apples to apples,” accounting for differences among NICUs in case mix. Examples of areas in which Network research has led to better reporting are mortality and LOS.

Figure 2 shows the standardized mortality ratio (SMR) for 191 NICUs in the Network database in 1996. The SMR is the ratio of the number of observed deaths at an NICU to the number of deaths predicted based on the characteristics of the patients treated at the NICU. Values >1 indicate that there were more deaths than would have been predicted based on the characteristics of the patients, whereas values <1 indicate that there were fewer deaths than would have been predicted.

The number of predicted deaths for each NICU is calculated using a multivariate logistic regression model that accounts for differences among NICUs in the types of patients they treat. The predictor variables are gestational age (weeks); gestational age squared (weeks²); major birth defect (yes, no); size for gestational age (<10th percentile, ≥10th percentile); multiple birth (yes, no); 1-minute Apgar score (1 to 10); gender; race/ethnicity (black, Hispanic, white, other); location of birth (inborn, outborn); and delivery by cesarean section (yes, no). All these variables are based on factors occurring before or immediately after birth and are not influenced by treatments received in the NICU. The model performs and fits the data well (area under the ROC curve 0.88; Hosmer–Lemeshow goodness of fit statistic with 8 degrees of freedom, 4.70; $P = 0.79$).

Two features of the data in Figure 2 are worthy of note. First, after accounting for differences in case mix, some units have significantly fewer deaths than expected, whereas others have significantly more deaths than expected. Previous analyses of Network data using similar models have shown that there is more variation in mortality rates among units than can be explained by differences in case mix or chance. Of course, it is possible that some of the unexplained variation could be attributable to risk factors for which we have not adjusted. Second, as indicated by the wide CIs, the estimates of the SMR for individual NICUs are relatively imprecise. This is because the number of very low birth weight infants treated at an individual NICU in a given year is small. The median number of infants weighing 501 to 1500 g treated at Network centers is approximately 80 per year. Estimates of the SMR based on samples of that size will always be relatively imprecise.

Figure 3 shows the risk adjusted total LOS for survivors at the 177 North American NICUs in the Network in 1996. The risk adjusted total LOS is the geometric mean of the total LOS before discharge home (including stays at hospitals to which infants were transferred before going home) calculated using analysis of covariance. This measure can be interpreted as estimating what the mean total LOS would have been had the individual NICU treated a group of infants with risk factors similar to those for the Network population as a whole.

The predictor variables used to estimate total LOS are birth weight (grams); assisted ventilation (IMV or high frequency: yes, no); respiratory distress syndrome (yes, no); bacterial sepsis on or before day 3 (yes, no); major surgery (yes, no); 1-minute Apgar score (1 to 10); size for gestational age (<10th percentile, ≥10th percentile); transfer status (not transferred, transferred for surgery, transferred other); delivery by cesarean section (yes, no); location of birth (inborn, outborn); gender; race/ethnicity

![Fig 2. The SMR at 191 NICUs participating in the Vermont Oxford Network in 1996. The SMR is the ratio of the number of predicted deaths to the number of observed deaths for infants weighing 501 to 1500 g. The number of predicted deaths is calculated for each NICU using a logistic regression model that includes terms for gestational age (weeks); gestational age squared (weeks²); major birth defects (yes, no); size for gestational age (<10th percentile, ≥10th percentile); multiple birth (yes, no); 1-minute Apgar score (1 to 10); gender; race/ethnicity (black, Hispanic, white, other); location of birth (inborn, outborn); and cesarean section (yes, no). The 95% CI for the estimated SMR is shown as a vertical bar. (Reprinted with the permission of the Vermont Oxford Network.]

![Fig 3. The adjusted average total LOS (Total LOS) in days for surviving infants weighing 501 to 1500 g at the 177 North American NICUs participating in the Vermont Oxford Network in 1996. Total LOS includes the initial NICU stay plus any additional days spent at other hospitals after transfer and before discharge home. The adjusted total LOS was calculated using analysis of covariance. The model included terms for birth weight (grams); assisted ventilation (yes, no); RDS (yes, no); early bacterial sepsis (on or before day 3); major surgery (yes, no); 1-minute Apgar score (1 to 10); size for gestational age (<10th percentile, ≥10th percentile); transfer status (not transferred, transferred for surgery, transferred other); delivery by cesarean section (yes, no); location of birth (inborn, outborn); gender; race/ethnicity]
(black, Hispanic, white, other); and major birth defect (yes, no). The analysis including hospital as a covariate had an R² of 0.73.

Similar to the mortality data, the LOS data also demonstrate dramatic variation among NICUs after adjusting for differences in case mix. The adjusted total stays range from <40 days to >75 days. We do not understand why this tremendous variation exists, but we intend that by reporting these data to the NICUs, they will begin to identify opportunities for reducing LOS where appropriate. As with the SMR data, it is possible that some of the unexplained variation in LOS is attributable to risk factors that we have not measured.

The data for mortality and LOS are examples of the widespread variations in practice and outcomes that are observed at Network NICUs. Similar variation can be seen for most of the outcomes and practices that we monitor. This suggests that there are substantial opportunities for improvement yet to be achieved.

QUALITY IMPROVEMENT

The database forms the cornerstone of the Network’s continuous quality-improvement efforts. The first step in using the Network database for this purpose is transforming the data into useful information. The database includes several million individual data items. These must be analyzed and the results synthesized and reported in a way that members can use to understand and assess their performance easily. Network reports are designed with these principles in mind.

Participating hospitals receive confidential quarterly and annual reports using the database to document their performance, identify trends over time, and compare performance to the Network as a whole. Reports include data on birth weight-specific incidence rates, as well as risk adjusted mortality, morbidity, and LOS. Birth weight strata provide homogeneous risk categories for reporting on morbidity and practice interventions. Network reports to members include figures similar to those for mortality and LOS shown above, which allow them to identify their own individual NICU. Of course, they cannot identify any of the other NICUs. This enables units to compare their risk-adjusted mortality and LOS with those in the rest of the Network.

A new reporting feature will be added to the Network reports in 1998. In addition to comparing units with the Network as a whole, we also will compare them with subgroups of NICUs providing a level of medical and surgical services similar to their own. Development of the system for classifying NICUs into these subgroups is the subject of ongoing research in the Network. This is an example of applied outcomes research at work and demonstrates further the relationship between research and quality-improvement activities in the Network.

Feedback of comparative, risk-adjusted practice and outcome information is crucial to supporting continuous quality-improvement efforts. However, although information is necessary, by itself it is not sufficient to foster lasting improvement. The information must be interpreted; opportunities for improvement identified; and appropriate change concepts developed, implemented, monitored, and maintained. In other words, information must be transformed into action. This requires knowledge, tools, skills, and resources for improvement. The Vermont Oxford Network is sponsoring an ongoing program of quality-improvement initiatives to help members achieve the necessary knowledge and skills and to develop the relevant tools. These initiatives include the Neonatal Intensive Care Collaborative Quality (NIC/Q) Project, and the Vermont Oxford Network Evidence-based Quality Improvement Collaborative for Neonatology, otherwise known as NIC/Q 2000.

THE NIC/Q PROJECT

During the past 3 years, we have applied a team approach to health care improvement with the goal of improving the effectiveness and efficiency of neonatal intensive care. The NIC/Q Project uses a collaborative model of quality improvement and benchmarking that was developed originally by industry and is now being applied successfully to health care.26–30

The major components of the project are 1) multidisciplinary collaboration within and among hospitals; 2) feedback of information from the Network database regarding clinical practices and patient outcomes; 3) training in quality-improvement methods; 4) site visits to project NICUs; 5) benchmarking visits to superior performers within the Network; 6) identification and implementation of “potentially better practices”; and 7) evaluation of the results. The project, involving 10 Network NICUs, has been funded by grants from the Center for the Future of Children of the David and Lucile Packard Foundation, and has been performed in collaboration with the Rand Corporation. The participating institutions in the NIC/Q Project are listed in Appendix I.

Since January 1995, teams from these hospitals have worked together in cross-institutional improvement groups and participated in an intensive series of large group meetings, site visits, and conference calls. They have chosen quality indicators, formed subgroups related to specific indicators, performed detailed process analyses and literature reviews, and participated in site visits to each other as well as to superior performers identified using the Network database. Based on these activities, the subgroups have developed a series of “potentially better practices,” which now are being implemented. The database is being used to monitor their impact.

The initial improvement goals chosen by the NIC/Q Project sites were a reduction in nosocomial bacterial infection for infants weighing 501 to 1500 g and reduction in chronic lung disease or death for infants weighing 501 to 1000 g. Chronic lung disease is defined as oxygen supplementation at 36 weeks’ postconceptional age. Six NICUs focused on reducing nosocomial infection; four units focused on reducing chronic lung disease.

Preliminary analyses have demonstrated significant improvement in both outcomes between 1994,
the year before the beginning of the project, and 1996, the year after implementation of the “potentially better practices” had begun. The overall rate of nosocomial infection at the six NICUs in the infection subgroup declined from 26.3% in 1994 to 20.9% in 1996 (P = 0.007). The rate of supplemental oxygen administration at 36 weeks for infants weighing 501 to 1000 g decreased from 43.5% in 1994 to 31.5% in 1996 (P = 0.03) at the four NICUs in the chronic lung disease subgroup. There was significant variation among NICUs with respect to the whether improvement occurred and the magnitude of improvement achieved within both subgroups. The improvements observed at the NICUs in these subgroups as a whole were significantly larger than were the changes observed at the 66 other Vermont Oxford Network centers in North America that participated in the Network, but not in the NIC/Q Project, from 1994 to 1996. A full report of the NIC/Q Project results is in preparation.

We are now building on the results of the original NIC/Q Project to create “The Vermont Oxford Network Evidence-based Quality Improvement Collaborative for Neonatology.” This project will run for a 2-year period, ending in the year 2000. It will be called NIC/Q 2000 to acknowledge its roots in the approaches developed during the first NIC/Q Project.

NIC/Q 2000

Our goal in the NIC/Q 2000 project is to create and evaluate an Evidence-based Quality Improvement Collaborative for Neonatology. The specific aims of the collaborative are:

1. to make measurable improvements in the quality and cost of neonatal intensive care,
2. to develop new knowledge, tools, and resources for quality improvement in neonatology, and
3. to disseminate the knowledge, tools and resources to the professional community.

The ideas and materials developed during NIC/Q 2000 will contribute to an evolving knowledge bank maintained by the Vermont Oxford Network (Fig 4). This knowledge bank will include clinical information; a new archive of clinical, organizational, and operational “better practices” and change concepts; a set of tools for improvement studies; and simple data collection instruments useful in the quality-improvement process. Our goal is to make this resource widely available to all interested in improving the quality of medical care for newborn infants.

The collaborative will assist multidisciplinary teams from 34 participating Network hospitals develop four key habits for clinical improvement: the habit for change, the habit for clinical practice as a process, the habit for collaborative learning, and the habit for evidence-based practice. The four key habits were conceptualized by Paul Plesk, a recognized leader in the field of quality improvement in health care, who served as the facilitator and consultant to the original NIC/Q Project and will serve a similar role in NIC/Q 2000. Multidisciplinary teams from participating hospitals will apply the four key habits to the continuous improvement of neonatal intensive care.

THE HABIT FOR CHANGE

The habit for change does not come naturally to individuals or organizations, yet it is the critical foundation on which improvement efforts must be built. The habit for change includes several components: identifying local leaders and champions for change; preparing the organization for change and fostering the specific skills, tools, and methods necessary to develop change ideas; and to test and implement them successfully.

There must be a sense of urgency for change. Individuals must understand why change is necessary. This requires strong internal champions for change who have a vision of what improvement can achieve and the skills to communicate that vision to everyone involved. Without leadership, change that results in significant improvement will not occur. In the NIC/Q 2000 collaborative, each organization will form a core leadership team consisting of two to four key individuals in the NICU who will have the overall responsibility for preparing their organization for change. The collaborative will assist the members of the leadership teams in developing the necessary skills.

The collaborative will use a new organizational assessment survey as a way of helping teams to understand their organization’s unique culture and to prepare for change. This survey tool, specific to the NICU environment, is being developed and tested in conjunction with Ross Baker, PhD, an expert in organizational assessment and quality improvement in health care. The results of the organization survey will be shared with the members of the collaborative so that they can understand how their organization compares with others in the collaborative from the perspective of organizational development. Throughout the collaborative, participants will work together to identify organizational and mana-
gerial “better practices” that create an organizational environment receptive to change and contribute to accelerated improvement.

The final component of the habit for change involves using a simple but proven model for accelerated improvement. The model, originally developed by Langley and colleagues, uses a “trial and learning” approach to improvement. Teams first must answer three simple questions:

1. What are we trying to accomplish?
2. How will we know that a change is an improvement?
3. What changes can we make that will result in an improvement?

A series of Plan-Do-Study-Act (PDSA) cycles then are used to test and implement the changes. In each individual cycle, specific plans are developed (Plan), the plans are conducted (Do), the results are evaluated (Study), and actions are taken based on what has been learned (Act).

As simple as these steps seem, organizations need training and assistance to integrate them into their routine behavior. The Breakthrough Series of the Institute for Health Care Improvement, under the leadership of Donald Berwick, has applied this model successfully in a variety of health care settings. The NIC/Q 2000 collaborative will apply the model to neonatal intensive care.

THE HABIT FOR CLINICAL PRACTICE AS PROCESS

Productive work is accomplished through processes. Yet it is a relatively recent development for physicians and other health care providers to view medical care as a process. In the past, physicians saw patient care strictly in terms of their own responsibility for history-taking, examination, diagnosis, and treatment. Other providers, including nurses, had similarly restricted views. A major advance attributable to the introduction of the concepts of continuous quality improvement into health care is the realization that patient care involves a complex system of interacting components and processes. Real improvement in quality requires an understanding of the total system and how its component parts are interrelated.

There is no doubt that the NICU is a complex system. The care of critically ill, very low birth weight infants requires a large multidisciplinary team and the application and coordination of a broad array of constantly changing technologies. NICU care must be integrated into the complex environment of a modern health care organization with all the competing demands for staffing, scheduling, financing, and other resources. A major goal of the proposed collaborative is to assist participants in understanding the processes involved in the NICU system. This will be accomplished in a number of ways.

Teams will be taught to describe and think about work in the NICU in process-oriented terms when appropriate: customers, suppliers, hand-offs, bottlenecks, sequence, flow, rework, and so forth. They will learn to be explicit about the care process through the use of flowcharts, care maps, and other standard documentation tools. Systems thinking will be stressed in a series of large and small group exercises and teaching sessions, using specific examples from participating NICUs. Team members will be helped to realize that processes within a system are interconnected and that when one is changed, others also may be influenced in a chain of intended and unintended consequences.

The crucial importance of measurement, to monitor process performance in an ongoing way, will be taught and supported. The database will be the basis for many of the measurements. Additional, simple measurement tools also will be needed, especially with regards to variations in the process of care. The group will receive instruction in the efficient collection, presentation, and interpretation of measurement data. The data will be used to identify areas of significant practice and outcomes variation. These will include clinical practices and outcomes; organizational factors such as leadership, communication, and unit culture; and operational issues related to the efficiency of the care process. Teams will be assisted in distinguishing appropriate variation from inappropriate, wasteful, or unintended variation. Throughout the collaborative, we will foster the habit for clinical practice as process through education, discussion, and the analysis of real-world NICU examples.

THE HABIT FOR COLLABORATIVE LEARNING

Collaborative learning is another critical skill that requires training and practice. The NIC/Q 2000 project will assist individuals and organizations to develop this key habit. Collaboration will mean creating close working relationships both among the disciplines within an institution as well as among the teams from different institutions. Physicians, nurses, administrators, and allied health personnel all must learn that individuals from other disciplines have knowledge and skills to contribute. The old autocratic model in which the physician gives “orders” and others follow will not work. New attitudes and behaviors are required. The working relationships among team members clearly changed over the course of the first NIC/Q Project as new patterns of interaction emerged and individuals developed new roles in the group. The NIC/Q 2000 project will foster this same type of multidisciplinary collaboration.

Cross-institutional collaboration has been used successfully in a number of different health care settings. The Northern New England Cardiovascular Disease Study Group has shown that collaboration and site-visiting can be important components of the improvement effort. O’Connor has reported that postsurgical mortality for coronary artery bypass graft procedures was reduced at a group of Northern New England hospitals that participated in a program involving performance feedback, quality-improvement training, and site-visiting. A key element of the program was intensive interinstitutional collaboration among multidisciplinary teams from
the different hospitals. During site visits, the teams were able to observe the routines of daily care in extreme detail. The lessons learned from these site visits lead to substantive changes in the processes and organization of care that were temporally associated with the decrease in mortality.

Benchmarking has been defined as the search for best practices that lead to superior performance. This type of benchmarking, coupled with in-depth process analysis and literature review, was used in the first NIC/Q Project to identify “potentially better” clinical practices. Multidisciplinary teams in the NIC/Q 2000 collaborative also will participate in benchmarking visits to Network hospitals, not necessarily in the collaborative, that have been identified using the database to achieve superior performance with respect to selected indicators. Questions remain about how to translate observations made in benchmarking visits into changes in clinical practice, particularly in cases where scientific evidence is lacking or weak. The knowledge gained during the NIC/Q 2000 collaborative will help answer these questions.

A theoretic perspective on benchmarking may be gained by considering the analogy with the role of recombination in biologic evolution. Just as recombination is an effective strategy for genetic evolution on rugged fitness landscapes, it also may represent an effective search strategy for other complex adaptive systems such as evolving medical technologies.

A critical element of successful cross-institutional collaboration is the open sharing of information about the processes and outcomes of care. There often is reluctance to share such information out of fear that the results could become public or be misused by others to gain a competitive advantage. The institutions participating in the NIC/Q Project adopted detailed guidelines regarding sharing of information, confidentiality, and publications. Similar guidelines will be developed for the proposed collaborative. Agreement with the principle of information-sharing within the collaborative subject to appropriate guarantees of confidentiality will be a prerequisite for participation in the project.

**THE HABIT FOR EVIDENCE-BASED PRACTICE**

Evidence-based medicine has been defined as “the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients.” The practice of evidence-based medicine requires self-directed learning and the acquisition and application of specific skills. Every patient encounter creates the need for information about diagnosis, prognosis, and therapy that then must be applied to make specific decisions that take into account the beliefs and values of the patient and family. In their recent book *Evidence-Based Medicine, How to Practice and Teach EBM,* David Sackett and colleagues have outlined five steps involved in the practice of evidence-based medicine. These are:

1. form clinical questions so that they can be answered,
2. search for the best external evidence,
3. critically appraise the evidence for its validity and importance,
4. actually apply the evidence in clinical practice, and
5. evaluate your performance as a practitioner of EBM.

We plan to incorporate teaching of these steps into the mission of the collaborative as we focus on assisting participants in developing the habit for evidence-based practice. This will be accomplished in several ways.

Each meeting of the collaborative will include formal teaching in the principles of evidence-based medicine using relevant examples from neonatology. The book by Sackett and colleagues will serve as an informal text. Sauve and associates have described the critically appraised topic (CAT), a formal, one-page summary in standard format that records the results of a critical appraisal of the literature. Resources related to the creation of CATs are available from the Center for Evidence-Based Medicine on the World Wide Web (http://cebm.jr2.ox.ac.uk/).

We plan to teach participants in the collaborative how to prepare CATs and encourage each team to submit CATs on a regular basis. These CATs can be cataloged, indexed, and made available to participants in printed or electronic form. This will encourage evidence-based thinking in daily practice, avoid duplication of effort among teams as they search for and evaluate the evidence, and focus attention on areas where formal overviews or randomized trials are needed.

The collaborative will reinforce the critical appraisal of the strength of health care recommendations and the quality of evidence on which the recommendations are based. “Better practice” concepts will be evaluated in relation to the published evidence. Participants in NIC/Q 2000 will use formal assessment tools to evaluate the evidence as they develop the habit for evidence-based practice.

Several approaches to grading the quality of evidence and the strength of the recommendations have been proposed. The original classification system was developed by the Canadian Task Force on the Periodic Health Examination and later adapted by the US Preventive Services Task Force. This rating system was used by the National Institutes of Health Consensus Panel in formulating and presenting its recommendations for antenatal corticosteroid therapy. Participants in the NIC/Q Project used a modification of this system for assessing the evidence supporting their “potentially better practices.”

There have been several advances in translating evidence from original studies into clinical recommendations. The major advance is the more widespread use of rigorous procedures and formal statistical techniques for combining the results from multiple studies in systematic overviews or meta-analyses. To reflect this development, the National Health and Medical Research Council of Australia has added a category to the rating system for the quality of evidence to account for systematic overviews of multiple trials.
Guyatt and colleagues, for the Evidence-based Medicine Working Group, have gone further. They have developed a rating system for health care recommendations based on systematic overviews that integrates three elements: the strength of the evidence in the overview; the threshold or magnitude of effect at which the benefit exceeds the risk of the therapy; and the relationships between magnitude of effect, the precision of the estimate of that effect, and the threshold.47

Systematic overviews of neonatal therapies are available in printed and electronic formats.48,49 The Neonatal Review Group of the Cochrane Collaboration is building a growing archive of systematic reviews in neonatal medicine. These can be accessed on the World Wide Web at a site maintained by the NICHD (http://silk.nih.gov/silk/cochrane). Unfortunately, there is currently insufficient data concerning most clinical practices to support systematic overviews. Participants in the NIC/Q 2000 collaborative therefore will often have to rely on less conclusive evidence.

A major challenge for participants in the collaborative will be to identify “potentially better practices” when little high-quality evidence is available in the literature. This situation will occur frequently. Observations made on benchmarking visits must be tempered by the availability of scientific evidence. The NIC/Q centers confronted this problem in the case of nosocomial infection. The benchmarking visits to two superior performing sites with low infection rates suggested that early use of central venous lines to avoid skin punctures for blood drawing and intravenous infusions might be responsible for the low rates of infections observed at these sites. These observations led to the hypothesis that improved skin care and increased skin integrity might lead to reduced infection rates. In response, the Network has designed a randomized trial of skin care practices to address this issue and gather the necessary evidence.

The chronic lung disease group of the NIC/Q Project confronted a similar issue regarding early therapy with corticosteroids. A literature review and meta-analysis suggested that early steroid therapy reduced the risk of chronic lung disease in very low birth weight infants.50 However, the evidence was judged insufficient to support routine introduction of this therapy. The group, therefore, recommended participation in the Network randomized trial of early steroids, rather than of implementation of a particular practice. A secondary gain of the collaborative will be the identification of questions requiring scientific study and the initiation of the appropriate randomized trials and outcomes studies. The habit for evidence-based practice is a cornerstone of this approach.

PROJECT TIMELINE

Thirty-four institutions currently are planning to participate in the NIC/Q 2000 collaborative. Before the first face-to-face meeting of the collaborative in September 1998, these institutions will form their local project teams, communicate the goals of the project widely within their organizations, administer the new organizational assessment survey, create an inventory of local measurement and improvement resources, identify areas of excellence and opportunity, and experiment with one PDSA cycle of rapid improvement. After the initial meeting, the teams will work together closely for 2 years to identify, test, and implement changes intended to improve the quality of neonatal intensive care. By applying the four key habits for clinical improvement and contributing their improvement knowledge to the continuously evolving Network Knowledge Bank, participants in the NIC/Q 2000 collaborative will create a valuable improvement resource that can be disseminated to other Network members and the wider professional community.

CONCLUSION

The Vermont Oxford Network is committed to improving the quality and efficiency of neonatal intensive care through a coordinated program of research, education, and quality improvement. The Network provides the infrastructure for this program, but it is only through the voluntary efforts of health professionals at the member institutions that its goals can be realized. The Network will continue to support these efforts by providing improved tools and new knowledge for the practice of evidence-based neonatology and the continuous improvement of neonatal intensive care.

APPENDIX I. Participants in the NIC/Q Project

Children’s Hospital Medical Center of Akron, Akron, OH
Dartmouth–Hitchcock Medical Center, Lebanon, NH
Emanuel Children’s Hospital, Portland, OR
Fletcher Allen Health Care, Burlington, VT
Miami Valley Hospital, Dayton, OH
Milton S. Hershey Medical Center, Hershey, PA
Minneapolis Children’s Medical Center, Minneapolis, MN
Parkview Memorial Hospital, Fort Wayne, IN
St. Francis Medical Center, Peoria, IL
Wesley Medical Center, Wichita, KS

APPENDIX II. Vermont Oxford Network Staff

Gary Badger, MS, Statistician
Paula Beales, Data Processor
Joan Briggs, Data Editor
Joseph H. Carpenter, MS, Network Statistician
Susan Dyer, Administrative Secretary
Gail Ewell, Data Clerk
Jean Fitts, Data Processor
Jeffrey D. Horbar, MD, Executive Director, Database Director
Michael Kenney, MS, Statistician
Kathy Leahy, RN, NNIP, NIC/Q Projects Coordinator
Jerold F. Lucey, MD, President
Nancy Morse, Database Manager
David Mortenson, Programmer
Ellen Willhite, Data Editor
Ollie Rutherfurd, Programmer
Roger F. Soll, MD, Clinical Trials Director
A. Lynn Stillman, Network Administrator

ACKNOWLEDGMENT

I thank Joseph H. Carpenter, MS, for performing the statistical analyses for Figures 2 and 3; Michael Kenney, MS, for creating these two figures; and the Vermont Oxford Network Staff for their dedication to accuracy, timeliness, and quality.

REFERENCES


The Vermont Oxford Network: Evidence-Based Quality Improvement for Neonatology
Jeffrey D. Horbar
Pediatrics 1999;103;350
The Vermont Oxford Network: Evidence-Based Quality Improvement for Neonatology
Jeffrey D. Horbar
Pediatrics 1999;103;350

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://pediatrics.aappublications.org/content/103/Supplement_E1/350