William Meadow and John Lantos

Epidemiology and Ethics in the NICU

Abstract

No medical professional is obliged to provide futile care. To be useful, however, futility determinations must be prospective and accurate. We wondered how accurately the professionals who work in the neonatal intensive care unit (NICU) recognized futile medical care as they were providing it, day-to-day, infant-to-infant. To estimate the accuracy of futility prognostications, we prospectively surveyed doctors and nurses in a NICU on their assessment of whether babies would survive or die. We then determined the number of times professionals predicted that a baby would die, and noted the accuracy of these predictions.

Overall, 802 infants were admitted to the NICU during this time period. We studied the 254 patients who received mechanical ventilation on at least one hospital day. Of the 254 ventilated patients, 55 (22%) died and 199 (78%) survived. Twentysix (13%) of the surviving infants survived after at least one day characterized by at least one estimate of 'death'. Indeed, eight infants survived despite having at least one hospital day in which ALL respondents predicted death. Whereas all respondents predicted survival in 78% of all patient days, these predictions were correct in 92%. On the other hand, all respondents predicted death at three consecutive days in 3% of all patient days and they were right in 82%. The percentages of correct predictions were considerably lower for the remaining cases in which the predictions were less uniform.

It is concluded that many futility assessments in the NICU are inaccurate. If certainty about futility were the only criterion that can justify a decision to withhold or withdraw life-sustaining treatment in the NICU, these data would make such decisions virtually impossible. These data also suggest caution in legitimizing policies that allow physicians to unilaterally determine that treatment will be futile. There is no quick and easy technical solution to the problems of prognostication.

Many people have an idea of what might be considered a 'good death.' For most people, a 'good death' is not one that takes place alone in an intensive care unit, tethered to high-tech life-support equipment, cared for by professionals who are unsuccessfully trying to prolong one's life. Instead, the good death takes place peacefully, surrounded by friends and loved ones, with careful attention to palliation of pain and suffering. By this view, each death in the Intensive Care Unit (ICU) can be interpreted as a failure of prognostication because if we knew the patient was dying, we would

have moved him or her out of the ICU. Nevertheless, most Americans who die today die in ICUs or other inpatient hospital settings. In one study, only 16% of deaths occurred at home, while 51% occurred in hospitals.[1] Less than 10% of Medicare beneficiaries who die ever get referred to a hospice and most of those are referred within a month of their death.[2]

From the perspective of the critical care doctor, the problem is not straightforward. Many patients die in ICUs, but many others, who are at great risk to die, are successfully treated and survive. If doctors could accurately distinguish those who are going to die from those who are going to survive, they could provide life-sustaining treatments to those who would survive and palliative care to those whose death was inevitable. Consequently, physicians, ethicists, economists and policy makers all recognize an urgent need to refine prognostic ability and accuracy.

Questions about prognosis and clinical decision making can be addressed under two broad moral frameworks. The first focuses on patient autonomy and the belief that patients (or, in the case of children, their parents) are in the best position to determine what type of health care they want. The goal for doctors, under this framework, is to empower patients by giving them the information and the authority that they need in order to determine the course of their treatment. With regard to end-of-life care, the central article of faith underlying this approach is the belief that, since patients *want* 'good deaths' as outlined above, and since they are not getting them, the problem must be that they do not have the knowledge or the power to make the choices that would give them what they want. The other broad moral framework focuses on medical futility. By this view, the problem is not that patients are disempowered. Instead, it is that both doctors and patients generally want and choose continued life-sustaining medical treatment unless the treatment is futile. Therefore, the challenge is not one of procedural empowerment but of prognostic refinement. If we can learn better how to determine whether a treatment is futile then doctors and patients will both be willing to forego it.

It has been difficult to develop refined prognostic techniques. In general, the problems hover around two related but separable questions. First, *how certain* do we have to be that treatment will be futile in order to treat it as such. Any quantitative assessment of prognosis will always have some uncertainty, some statistically definable 'confidence interval' around a point estimate. However, determining the sufficient degree of precision of the estimate will always require a value judgment. The second question concerns the *particular outcomes* that 'count' in the calculation of futility. Death is the easy one. The harder ones are whether any particularly dismal quality of life should also count as a treatment failure.

There are many ethical dilemmas in the Neonatal Intensive Care Unit (NICU), and almost as many solutions as dilemmas. Religion, philosophy, natural law, civil law, criminal law, to name but a few disciplines, have each been invoked as a source of authority to resolve the inevitable conflicts arising at the confluence of uncertain outcome, physical pain, and financial expenditure. This chapter primarily focuses upon an epidemiological research agenda for such dilemmas. The discussion will be divided into three parts: first, conclusions derived from retrospective studies of NICU mortality; next, conclusions derived from prospective

studies of NICU mortality; and finally, proposals for prospective studies of NICU morbidity.

Retrospective insights into futility

Who dies in the NICU and when do they die? Two different populations of NICU babies raise very different moral issues. One sub-population at high risk of mortality is the group of babies with severe congenital anomalies. The other high-risk population consists of extremely premature or low birthweight babies. Our studies have focused on the second populations: very premature, extremely low birthweight (ELBW) babies. The ethical dilemmas raised by these two populations are quite distinct. For babies with congenital anomalies, the prognosis is usually fairly well defined and understood. For complex congenital heart disease, for example, the mortality rates with surgery are well defined.[3] For myelomeningocele, the long-term morbidity has not changed much in twenty years.[4] Because the prognosis for these babies is relatively clear, the dilemmas focus on whether the burdens of treatment outweigh the benefits.

The dilemmas for ELBW babies are different. For them, the range of outcomes is enormous, from death or neurologic devastation to completely intact survival. Furthermore, outcomes have changed so dramatically over the past twenty years that predicting long term outcomes today is tenuous. These ELBW infants account for the vast majority of deaths in the NICU.

The majority of larger infants who die succumb to congenital anomalies. Many of these deaths are post-neonatal. At present, in industrialized countries, babies of less than 500 gram birthweight rarely survive. Above 1000 gram, survival rates are higher than 90%. Consequently, virtually all of the ethical controversy in the NICU focuses on babies between 500 and 1000 gram birthweight. This corresponds roughly to between 24 and 28 weeks of gestational age.

More interesting than which patients die is when they die. The vast majority of doomed infants die quickly. The median day of death in this NICU population is roughly the third day after admission. Across many NICUs with many varied practice styles, this phenomenon is remarkably robust. In almost all reports, the large majority of doomed NICU infants die early, and the smallest babies, who are at the greatest risk of dying, die the soonest.[5]

Two conclusions follow directly from these observations: one with profound implications for individual infants, the second with implications for public policy. Consider a group of infants born at 500 to 600 gram on their first day of life (DOL). Overall, only one infant in four in this group will survive. However, consider the same population three days later. Most of the doomed infants have now died, leaving a markedly different prognosis for the residual population of DOL 4 survivors. Even the tiniest infants who survive to DOL 4 have a very reasonable (over 70%) likelihood of surviving to discharge. Thus, although birthweight is a powerful predictor of survival on DOL 1, birthweight carries much less prognostic significance only a few days later.

The second, perhaps less obvious, conclusion that derives from these observations is that the relative proportion of medical resources expended on doomed ELBW infants does *not* depend either on birthweight or mortality risk. Rather resources expended on doomed infants remain consistently low across all birth weight groups. This is true because, although smaller babies are more likely to die, they also tend to die after far shorter hospital stays. Consequently, although more 600 gram infants die than 900 gram infants, they die earlier and consume fewer medical resources during their brief lives. Furthermore, the few 600 gram birth weight infants who do survive stay in the NICU a long time before discharge (approximately 100 days). Consequently, considering the 600 gram cohort as a whole, many more bed days are allocated to surviving infants than doomed ones, despite the fact that there are many more doomed infants than survivors, precisely because the doomed babies stay so much shorter than the survivors. Regardless of birthweight, roughly 85% of bed-days (equivalent to 85 cents of every NICU dollar) are allocated to infants who will be discharged alive.[6]

Prospective insights into mortality

No medical professional is obliged to provide futile care. To be useful, however, futility determinations must be prospective and accurate. Physicians and other medical caretakers often have intuitions about the likelihood of survival for patients in their care. Previous studies suggest that intuitions of survival garnered on the day of admission to an ICU or NICU correlate significantly (in a statistical sense) with patient outcomes.[7,8] There are, however, two problems with these observations. First, the correlations are not strong; that is, there is a lot of slippage between predictions of non-survival and actual death. Second, predictions on the day of admission do not take into account the 'trial of therapy' that is inherent in ICU care. No one is admitted to an ICU for 'hospice' care. Rather, ICU patients get aggressive, high-tech care in an attempt to prolong their lives. Patients, it is often said, 'declare themselves' in response to their therapy, but these declarations may take time before they are interpretable. Consequently, instead of analyzing one-time predictions on the day of admission for ICU patients, a more ethically relevant approach might be to analyze serial assessments made daily for the same ICU patient.

We wondered how accurately the professionals who work in the NICU recognized futile medical care as they were providing it, day-to-day, infant-to-infant. To find out, we asked doctors and nurses in our NICU one single question every day about patients in their care: do you think this child will die before hospital discharge, or live to go home to his family? We obtained responses from multiple caretakers for each infant for each day.

Overall, 802 infants were admitted to the NICU during this time period. We studied the 254 patients who received mechanical ventilation on at least one hospital day (infants treated with nasopharyngeal continuous positive airway pressure were excluded from this analysis). For each ventilated patient, on each day, nurses (both primary nurse and other nurses 'covering' the patient in the NICU), residents, fellows,

and attendings were approached and asked 'Do you think this child is going to live to go home to his family, or die before hospital discharge?' In addition to 'live' or 'die', each respondent was allowed to answer 'uncertain' if she could not comfortably predict either survival or non-survival for that infant, on that day.

Of the 254 ventilated patients, 55 (22%) died and 199 (78%) survived. Not surprisingly, the non-survivors were on average smaller and had shorter gestations than the surviving infants. Almost half of the non-survivors (27/55) were less than 750 gram at birth. In contrast, nearly three-fourths (148/199) of the survivors weighed over 1000 gram at birth.

Prediction profiles were obtained for 230 of the 254 (91%) infants who received mechanical ventilation during the study period: 192/199 (96%) surviving infants and 38/55 (69%) non-surviving infants. All of the non-survivors who were not profiled died in the first 72 hours of life, except two infants who were born and died during a one-week scheduling interruption. The 230 patient profiles contain predictions obtained on 2867 patient days. The average number of daily predictions for each ventilated infant was four. Consequently, approximately 11,000 predictions of patient outcomes were compiled during the 48 weeks of this study. There was no significant difference in the number of daily predictions obtained for non-survivors versus survivors.

Prediction profiles for non-survivors

Death between DOL 1-3: 21 of the 55 (38%) non-survivors died in the first three days of life. The median day of death for these infants was DOL 2. Six of these infants received prediction profiles, all of which reflected uniform prediction of death by every health care provider on every day. Each of the other infants in this category died before any outcome predictions were obtained.

Death between DOL 4-10: 12 of the 55 (22%) non-survivors died between DOL 4 and 10. Ten of these doomed infants received prediction profiles. As a group, these profiles were also homogeneously both dismal and accurate. Seven of the ten infants in this category had 100% prediction of death on every DOL from birth to the day of death. For two other patients in this group the profile differed only slightly: on at least one day during the first 72 hours of life survival was thought likely by at least one respondent. However, by DOL 4 non-survival was uniformly and accurately predicted. Thus for nine of the ten non-surviving infants in this group, on each day between DOL 4 and DOL 10, no respondent thought that the child would survive (that is, the prediction of survival to discharge was 0% for each hospital day).

Death after DOL 10: 22 of the 55 (40%) non-survivors died after DOL 10. All of these infants received prediction profiles. In contrast to the homogeneity that characterized profiles of infants who died before DOL 10, the 22 later-dying infants were a heterogeneous group. Only five (22%) of these 22 late-dying infants had the uniform prediction of death that categorized predictions for infants who died prior to DOL 10. Each of the other seventeen late-dying infants was predicted to live by many (if not all) observers on many (if not all) hospital days. Eleven (50%) of these late-dying infants suffered, with little warning, a fatal medical catastrophe (NEC, sepsis, pneu-

monia, et cetera). The rapid and unexpected nature of their demise is emphasized by the observation that for seven of these late-dying infants, not even one day of their hospital stay was marked by 0% prediction of survival. Six (27%) of 22 late-dying infants had prediction profiles categorized by considerable uncertainty, both within respondents and across days. That is, several hospital days were characterized by 'pessimism' (that is, low predictions of survival), alternating with periods of 'optimism', characterized at times by up to 100% prediction of survival. These infants often survived for many weeks prior to their death.

Prediction profiles for survivors

Prediction profiles for survivors reflected two distinct hospital courses. The vast majority of surviving infants were predicted by all (or almost all) observers to survive on all (or almost all) days of mechanical ventilation. One hundred fiftyseven (81%) of the 193 survivors had this consistent, accurate prediction profile and for 136 (70%) of 193 surviving infants, every NICU ventilator-day was characterized by 100% prediction of survival. Twentyone other survivors had profiles nearly as positive: for these infants, a brief period of uncertainty was followed by increasing confidence in the likelihood of survival, but at least 90% of their hospital ventilator-days were characterized by 100% prediction of survival. At the other end of the continuum, 26 (13%) of 193 surviving infants survived somewhat unexpectedly; that is, after at least one day characterized by at least one estimate of 'death'. Indeed, eight infants survived despite having at least one hospital day in which *all* respondents predicted death.

Accuracy of predictions of survival and non-survival

Predictions of survival for ventilated infants were very common and very accurate. Over three quarters of NICU days occupied by ventilated patients were characterized by uniform prediction of survival. Of these predictions, 92% were correct. Non-survival predictions were much less common and much less accurate. The more people who consistently predicted non-survival, the more accurate the predictions were. However, even when every health care professional predicted that a baby would die for three days in a row, they were wrong 18% of the time. The percentage of accurate prognostications is shown in the Table.

Table. Percentage of positive and negative prognostications.

Prognostication	% of patient days	% correct
Uniform prediction of survival	78	92
One prediction of death	18	40
50% prediction of death	11	51
100% prediction of death during one day	5	69
100% prediction of death during three days	3	82

Conclusions and future research questions

The study has methodologic limitations. First, our data may reflect a self-fulfilling prophecy; that is, once 'non-survival' is predicted is the balance of NICU care 'tilted' to produce the demise of the infant? We saw no evidence of such behavior during the study period. In fact, the overwhelming majority of non-survivors in our study did not have DNR orders. Second, our predictions may reflect a 'herd' phenomenon; that is, once the opinion of 'non-survival' was articulated (particularly by the attending physician), did others 'jump on the bandwagon'? This possibility is difficult to evaluate, as the opinions may equally have reflected shifts in the bodies of the infants as in the minds of the evaluators. Nevertheless, we explored this possibility in a pilot study by comparing predictions of our respondents to predictions of experienced NICU nurses who did not participate in rounds or provide direct patient care during the study period. There was substantial agreement between our 'blinded' respondents and our study respondents.

These data carry a number of important implications for discussions about prognostications of medical futility and the withholding and withdrawing of life-sustaining treatment in the NICU. First, they suggest that very little recognizably 'futile' care is being provided. That is, there were very few circumstances in which every professional agreed that the baby would not survive, treatment was extended, and the baby eventually died. To the extent that prolonged treatment was provided to babies who ultimately died, almost all of their deaths were unpredictable. Second, our data raise the disturbing possibility that many futility assessments are inaccurate. This raises some interesting problems. If medicine, like meteorology, is an inexact science, long-range predictions of death, at least in the NICU, may be as imperfect and as useless as long range weather forecasts. Furthermore, if certainty about futility were the only criterion that can justify a decision to withhold or withdraw life-sustaining treatment in the NICU, these data would make such decisions virtually impossible. We would suggest that there are situations in which withdrawal of care is appropriate, that such decisions are always based on probabilistic information about outcomes, and that certainty is therefore an impossible threshold and an illusory criterion for such decisions.

Finally, we have demonstrated that the 'distributive justice' argument strongly favors continued NICU care. The vast majority of NICU resources are directed to infants who ultimately survive to go home to their families, tenfold more than ICU resources directed toward sick adults.

We have only begun to explore the implications of predictions of morbidity. Future research should focus on the relationship between predictions of mortality and ultimate outcomes for patients who survive. A great deal of population-based literature recounts the likelihood of morbid outcomes (almost always a combination of motor spasticity and cognitive impairment) as a function of risk factors for NICU patients. In brief, this work shows that the higher the risk of death, the higher the risk of survival with impairment.[9]

However, just as with mortality, morbidity is more importantly described from a prospective viewpoint. Clinicians are not faced with 'a population' (although public

policy makers are). Rather doctors and nurses deal with patients one at a time. What do we know about the accuracy of predictions of impairment in survivors of NICU care while the infants require life-support in the NICU, as opposed to at their two-year check-up? The answer, in short, is very little. Although some physiologic events in the NICU have clearly been correlated with subsequent impairment, very little attention has been paid to caretaker intuitions (either serial or 'one-time') about morbid outcomes.

One could design a prospective morbidity study closely paralleling the prospective mortality study described in the section above. One could ask caretakers on a daily basis whether the infant in their care was going to 'live but be impaired', with various degrees of impairment specified or not. Correlation of these intuitions with subsequent outcomes would provide at least a first-cut answer to the predictive value of intuitions of morbidity. It may turn out that the predictions of outcome by doctors are not that bad, but that mortality is not the only bad outcome to be avoided. Survival with severe neurological deficits may be as bad or worse in the minds of some parents.

Parental perception of the goals of NICU care is a second important area for future research. Do parents feel that they are adequately involved in decision-making now? For parents of babies who died, do they feel that they achieved a 'good death?' If not, what mechanisms might facilitate more truly shared decision-making? The goal should be to combine the best epidemiological data with the best methods of sharing that data to insure that parents understand, and then seek the best decision for each infant within the inevitable constraints of prognostic uncertainty.

References

1. Polissar L, Severson RK, Brown NK. Factors affecting place of death in Washington State, 1968-1981. J Community Health 1987;12:40-55.
2. Christakis NA, Escarce JJ. Survival of Medicare patients after enrollment in hospice programs. N Engl J Med 1996;335:172-178.
3. Samanek M, Voriskova M. Congenital heart disease among 815,569 children born between 1980 and 1990 and their 15-year survival: a prospective Bohemia survival study. Pediatr Cardiol 1999; 20:411-417.
4. Alexander MA, Steg NL. Myelomeningocele: comprehensive treatment. Arch Phys Med Rehabil 1989;70:637-641.
5. Meadow W, Reimshisel T, Lantos J. Birth weight-specific mortality for extremely low birth weight infants vanishes by four days of life: epidemiology and ethics in the neonatal intensive care unit. Pediatrics 1996;97:636-643.
6. Lantos J, Mokalla M, Meadow W. Resource allocation in neonatal and medical ICUs. Epidemiology and rationing at the extremes of life. Am J Resp Crit Care 1997;156:185-189.
7. Stevens SM, Richardson DK, Gray JE, et al. Estimating neonatal mortality risk: an analysis of clinicians' judgments. Pediatrics 1994;93:945-950.
8. Kruse J, Thill-Baharozian M, Carlson R. Comparison of clinical assessment with APACHE II for predicting mortality risk in patients admitted to a medical intensive care unit. JAMA 1988;260:1739-1742.
9. Pinto-Martin JA, Riolo S, Cnaan A. Cranial ultrasound prediction of disabling and non-disabling cerebral palsy at age 2 in a low birth weight population. Pediatrics 1995;95:249-254.