Prediction of adverse outcomes in children with sickle cell anemia: a study of the Dallas Newborn Cohort

Charles T. Quinn,1-3 Nancy J. Lee,2,3 Elizabeth P. Shull,2,3 Naveed Ahmad,3 Zora R. Rogers,1-3 and George R. Buchanan1-3

1Division of Hematology-Oncology, Department of Pediatrics, University of Texas Southwestern Medical Center at Dallas; 2Southwestern Comprehensive Sickle Cell Center, Dallas, TX; and 3Children’s Medical Center, Dallas, TX

Introduction

Sickle cell anemia (SS), the homozygous form of sickle cell disease (SCD), has protean manifestations that are mostly unpredictable.1 Therefore, it has been difficult to identify young children with SS who are at highest risk of adverse outcomes before their irreversible organ damage has occurred. A robust predictive model would allow early, tailored therapy to prevent adverse outcomes. Transcranial Doppler ultrasonography is an example of a successful screening tool that can identify children with SS at high risk of stroke.2 Transcranial Doppler screening in conjunction with directed chronic transfusion programs has begun to decrease the incidence of stroke.3 Besides stroke, the other serious complications of SCD are difficult to predict in individual patients.1 As such, clinicians often wait for children to manifest severe disease before initiating treatments, such as hydroxyurea, chronic transfusions, or stem-cell transplantation.

To address this challenge, the Cooperative Study of Sickle Cell Disease (CSSCD) developed a model to predict severe disease in its infant cohort.4 Severe disease was defined as one of 4 adverse outcomes: death, stroke, frequent pain, or recurrent acute chest syndrome (ACS). The probability of experiencing one of these adverse outcomes by 10 years of age could be predicted by a multivariable model that included the following predictors: dactylitis in the first year of life (“early dactylitis”), the mean steady-state hemoglobin (Hgb) concentration in the second year of life, and the mean steady-state leukocyte count in the second year of life. The CSSCD infant cohort accrued patients from 1978 to 1988, which was before both newborn screening and prophylactic penicillin were universal and when fatal bacterial sepsis was more common than it is now. Indeed, after stroke, death was the second most frequently predicted adverse outcome. The CSSCD early prediction model has not been validated in other cohorts, so we wished to test its performance in a contemporary SCD cohort.

The Dallas Newborn Cohort (DNC) is a large, single-center newborn inception cohort that is independent of the CSSCD.5 The rate of fatal bacterial sepsis in the DNC is low because of universal newborn screening, prophylactic penicillin, and recent use of the heptavalent conjugated pneumococcal vaccine. We previously showed that painful events, including dactylitis, in the first 3 years of life had limited prognostic utility for adverse outcomes in the DNC.6 However, unlike the CSSCD, we did not restrict the analysis of dactylitis to episodes in the first year of life, and we did not study laboratory markers.6 Therefore, we sought to test the performance of the multivariable CSSCD model. We hypothesized that it did not predict adverse outcomes in the DNC.

Methods

This is a retrospective cohort study in which the CSSCD early prediction model was applied to the DNC. The CSSCD model has not been used to direct therapy for members of the DNC.

Subjects

The DNC is a newborn inception cohort that includes patients with the 4 common SCD genotypes: SS, sickle-hemoglobin C disease (SC), sickle-β-1-thalassemia (SB1), and sickle-β-ß-thalassemia (SB2).5 It is noteworthy among SCD cohorts because universal newborn screening for hemoglobinopathies in Texas identified all members, and all subjects with...
SS or Sβ0 were prescribed prophylactic penicillin until at least age 5 years. Accrual began in 1983 and is ongoing. This analysis includes follow-up through February 9, 2005. Members are tracked prospectively in our center’s comprehensive SCD database. The Institutional Review Board of the University of Texas Southwestern approved the use of the clinical database for this project and waived the requirement for informed consent.

Inclusion and exclusion criteria

DNC subjects were included in this analysis if they had SS or Sβ0, were born on or after January 1, 1990, had their first visit to our center in the first year of life, and were at least 2 years of age at study closure. We studied individuals with SS or Sβ0 as a single group because of their known clinical similarity and the very small number of Sβ0 subjects. We originally aimed to analyze all eligible cohort members, beginning with the inception of the cohort in 1983. However, to best use our resources, we chose the cutoff of January 1, 1990, instead because we found that complete data were routinely available in our shadow charts and the hospital’s electronic medical record only for subjects born after this date. We chose this cutoff date during the data-collection phase but before we performed any analysis. Because leukocyte count and the steady-state Hgb in the second year of life were studied as predictors, we excluded those few patients who did not attend clinic at least once in steady-state or began chronic transfusions or hydroxyurea during the second year of life. Subjects who had incomplete records were also excluded.

Definitions and measurements

Early predictors. The same 3 predictors of adverse outcomes used in the CSSCD predictive model were studied here: (1) occurrence of dactylitis in the first year of life, (2) steady-state Hgb concentration in the second year of life, and (3) steady-state leukocyte count in the second year of life.4 Steady-state Hgb concentration and total leukocyte count were defined as the mean of all Hgb and leukocyte measurements obtained at routine visits during the second year of life when the child was experiencing no acute medical problems. We considered only blood counts that were known to be obtained 3 or more months after a transfusion. All DNC leukocyte counts were automatically corrected for the presence of nucleated red blood cells by the cell counter, unlike the CSSCD. Results of complete blood counts were obtained from the electronic and paper-based medical records. A steady-state Hgb less than 7 g/dL was classified as “severe anemia”; a steady-state leukocyte count greater than or equal to 20 000/mm3 was classified as “leukocytosis.”

The occurrence of dactylitis was ascertained by query of our database for “dactylitis” and permutations of “hand-foot syndrome.” We further manually reviewed all subjects’ medical records for documentation of dactylitis in their clinic visit notes. Our standard clinical practice is to question parents and families of young children about the occurrence of interim complications, including dactylitis, at each clinic visit and to record these findings in the medical record.

Adverse outcomes. The same 4 adverse outcomes predicted in the CSSCD infant cohort were studied here: (1) death of any cause; (2) clinically overt first stroke; (3) frequent pain (average of ≥ 2 painful events per year for the entire follow-up period, with ≥ 2 events per year for 3 consecutive years); and (4) recurrent ACS (average of 1 or more episode per year for the entire follow-up period with ≥ 1 episode per year for 3 consecutive years).4 However, we considered only hospitalizations for pain in this analysis, unlike the CSSCD, because we do not systematically track episodes of pain treated only at home or in outpatient facilities (eg, emergency rooms or clinics). Outcomes were identified by query of the database and review of selected medical records. If a subject experienced more than one adverse outcome, only the first adverse outcome was considered. Only episodes of pain or ACS that occurred before the start of any disease-modifying therapy (hydroxyurea, chronic transfusions, or stem cell transplantation) were included because such treatments usually alter their natural frequency. If the reason for any single hospitalization was recorded in the database as both a painful event and ACS (eg, painful event complicated by ACS), the hospitalization was counted as an episode of ACS only because ACS was considered to be the more severe complication. ACS is defined clinically in our center as an acute pulmonary illness that is characterized by a new radiographic pulmonary infiltrate and some combination of fever, hypoxemia, thoracic pain, and signs and symptoms of respiratory illness.7

Length of follow-up. We calculated the length of follow-up for each subject as the time from study entry (the first visit in our center) until the end of study (February 9, 2005) or, if the patient was lost to follow-up, until the date of last clinical visit in our center. Follow-up time was also truncated at the first occurrence of any of the following: death, clinically overt stroke, criteria being met for frequent pain or recurrent ACS, initiation of hydroxyurea or chronic transfusions, or stem cell transplantation.

Statistical analysis

Univariate analyses. We first examined the relationships between the 3 individual predictors and the composite of adverse outcomes. For early dactylitis and severe anemia, we used the Fisher exact test and odds ratios (OR). For leukocyte count we used a 2-sided t test for equality of means (equal variances not assumed) and calculated a mean difference and a 95% confidence interval (CI).

Multivariate analyses. We then calculated the expected probability of an adverse outcome for each DNC subject with the logistic regression equation used to predict adverse outcomes in the CSSCD infant cohort. This equation includes the following covariates: follow-up time; early dactylitis (yes or no); severe anemia (yes or no); and mean steady-state leukocyte count (continuous) in the second year of life. The parameter estimates for this predictive model were communicated by CSSCD investigators (S.T. Miller, SUNY Health Sciences Center, Brooklyn, NY and L.A. Sleeper, New England Research Institutes, Watertown, MA, oral communications, November 22, 2006). The predicted probabilities from this calculation were classified according to CSSCD definitions as “low risk” (probabilities < 0.09), “medium risk” (0.09-0.35), or “high risk” (≥ 0.36). We then compared the predicted and the observed outcomes in the DNC. A receiver-operating characteristic (ROC) curve was generated to describe the relationship between the calculated probability variable and the observed outcome. We also used the Cox proportional hazards method to model the time to occurrence of an adverse outcome while controlling for the following baseline covariates: early dactylitis (dichotomous), hemoglobin concentration (continuous), and total leukocyte count (continuous).

Data were analyzed using SPSS version 13.0 statistical software (SPSS, Chicago, IL). P values less than .05 were considered statistically significant. We did not adjust for multiple comparisons.

Results

Characteristics of subjects

We identified 206 subjects who met the inclusion criteria. Thirty-eight were ineligible for the following reasons: 14 had no clinical visits in the second year of life; 4 had no clinic visits while in steady state; 3 were treated as infants with BABY HUG study medicine, which may have been hydroxyurea; 2 were begun on chronic transfusions before 2 years of age; and 15 had incomplete medical records. The remaining 168 children provided 1188 patient-years of follow-up. Ninety-two were male and 76 were female; 163 had SS and 5 had Sβ0. Table 1 shows ages and hematologic data. Table 2 shows the frequency of the occurrence of the early predictors, both individually and in the combinations that conferred a high probability of an adverse outcome at 10 years of age in the CSSCD infant cohort.3 The 4 subjects who had high-risk combinations had a mean follow-up of 9.7 years (median, 12 years). Table 3 details the adverse events in the DNC. The mean time to an adverse event was 4.7 years (median, 3.7 years; interquartile range, 2.3-6.0 years).

Two subjects experienced 2 adverse events each. One had a stroke at 2 years of age, and he died of complications of recurrent strokes at 7 years of age. The second had a stroke at 6 years of age, and she also met the criteria for recurrent ACS at the time chronic
transfusions were begun. We considered only these 2 patients’ initial overt strokes for purposes of this analysis.

The frequency of adverse events among the 38 ineligible subjects was not different from the 168 eligible subjects (data not shown). We also found there to be neither a significant association (\(P = .182\)) between the time period of analysis and the rate of adverse outcomes nor a trend for a difference (\(P = .108\); data not shown), so it is reasonable to assume that the rate of adverse outcomes was constant throughout the study period.

Ten subjects were treated with hydroxyurea, of whom only 3 met the strict CSSCD criteria for an adverse event (frequent pain) before starting therapy. Fifteen were treated with chronic transfusions, of whom 14 met CSSCD criteria for an adverse event before starting therapy (13 had stroke; 1 had frequent pain). One subject underwent stem cell transplantation because of overt stroke.

### Performance of the model

**Univariate analyses.** The occurrence of early dactylitis was not associated with adverse outcomes: 13.8% of subjects with early dactylitis had an adverse outcome compared with 13.7% of those without dactylitis (\(P > .999\); OR = 1.01 [95% CI, 0.32-3.23]). The occurrence of severe anemia was also not associated with adverse outcomes: none of the subjects with severe anemia had an adverse outcome compared with 14.1% of those with without severe anemia (\(P > .999\); OR = not defined). Although not statistically significant, the direction of this association was the opposite of that observed in the CSSCD. A higher leukocyte count was associated with adverse outcomes: the mean leukocyte count in the second year of life was 16,560/\text{mm}^3 for those who had an adverse outcome compared with 13,950/\text{mm}^3 for those who did not (\(P = .045\), mean difference 2615 [95% CI, 61-5170]).

**Multivariate analyses.** The predicted probabilities of adverse outcomes for DNC subjects, calculated by the CSSCD logistic regression equation, ranged from 0.01 to 0.52 (mean, 0.11; median, 0.07). In the DNC, 57.1% were predicted to be “low risk,” 38.7% were “medium risk,” and 4.2% were “high risk.” This compares to a distribution in the CSSCD of 44%, 53%, and 3%, respectively. Fifteen (16.3%) “low risk” and 8 (12.3%) “medium risk” DNC subjects had adverse outcomes. None of the 7 DNC subjects who were classified as “high risk” actually experienced an adverse outcome as defined by the CSSCD criteria. One of these 7 subjects began treatment with hydroxyurea at 12.5 years of age for recurrent painful events. His prehydroxyurea rates of pain and ACS were 0.96 and 0.32 hospitalizations/year, respectively. The remaining 6 subjects with a mean follow-up of 13.1 years (median, 12.9 years) have never been treated with hydroxyurea, chronic transfusions, or stem cell transplantation.

Figure 1A depicts the ROC curve for the relationship between predicted probability and observed adverse outcomes. The area under the curve is less than 0.5 (0.409 [95% CI, 0.308-0.510; \(P = .161\)), so the performance of the model is not better than random prediction. From this analysis, a predicted probability of 0.36 yields a sensitivity of 0% and specificity 95%. This differs from the sensitivity of 23% and specificity of 91% reported in the CSSCD. Because leukocyte count was the only statistically significant predictor variable in the univariate analysis, we also performed an ROC analysis with leukocyte count as a single predictor (Figure 1B). The area under this ROC curve was 0.634 (95% CI, 0.517-0.752; \(P = .039\)). A total leukocyte count of approximately 20,000/\text{mm}^3 had a sensitivity of 21.7% with a specificity of 89.7%.

In the Cox proportional hazards model, the omnibus test of model coefficients was not significant (\(P = .068\)). Early dactylitis (\(P = .548\)), hemoglobin concentration (\(P = .140\)), and total leukocyte count (\(P = .089\)) were not significant predictors of the time to an adverse outcome.

### Discussion

We found that the CSSCD early prediction model was not better than random prediction in the DNC. Most subjects who experienced adverse events were predicted to be at low risk for adverse events. No subject who was predicted to be at high risk actually experienced an adverse outcome. One component of the CSSCD model, the total leukocyte count in the second year of life, was modestly associated with adverse outcomes in univariate analysis, but not in the Cox proportional hazards model. Leukocytes are known to be involved in the process of vaso-occlusion, and leukocytosis in adults is associated with an increased frequency of ACS and stroke. As such, early leukocytosis could also be an indicator of a patient’s tendency for more frequent vaso-occlusion throughout life. The 2 other components of the model, early dactylitis and hemoglobin concentration in the second year of life, were not predictive of outcomes in either the univariate or multivariate analyses. We conclude that the model is not clinically useful, at least in

### Table 1. Characteristics of DNC subjects

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Mean</th>
<th>Median</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ages and follow-up</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age at first visit, d</td>
<td>87.1</td>
<td>79</td>
<td>24-344</td>
</tr>
<tr>
<td>Age at last follow-up, y</td>
<td>8.2</td>
<td>8.0</td>
<td>2.0-15.1</td>
</tr>
<tr>
<td>Length of follow-up, y</td>
<td>7.1</td>
<td>6.5</td>
<td>1.2-14.8</td>
</tr>
<tr>
<td>Hematologic data in 2nd year of life</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blood counts per subject, no.*</td>
<td>2.7</td>
<td>3</td>
<td>1-6</td>
</tr>
<tr>
<td>Hemoglobin concentration, g/dL</td>
<td>8.8</td>
<td>8.7</td>
<td>6.1-13</td>
</tr>
<tr>
<td>Total leukocyte count, (\times 10^9/\text{mm}^3)</td>
<td>14.3</td>
<td>13.4</td>
<td>5.8-34.3</td>
</tr>
</tbody>
</table>

*Subjects with 1 blood count = 33; 2 counts = 42; 3 counts = 51; 4 counts = 31; 5 counts = 8; and 6 counts = 3.

### Table 2. DNC subjects who had “high risk” early predictors

<table>
<thead>
<tr>
<th>Early predictors</th>
<th>Subjects, N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual predictors</td>
<td></td>
</tr>
<tr>
<td>Early dactylitis*</td>
<td>29 (17.3)</td>
</tr>
<tr>
<td>Severe anemia†</td>
<td>5 (2.9)</td>
</tr>
<tr>
<td>Leukocytosis‡</td>
<td>20 (11.9)</td>
</tr>
<tr>
<td>“High risk” combinations§</td>
<td></td>
</tr>
<tr>
<td>Early dactylitis and severe anemia</td>
<td>0</td>
</tr>
<tr>
<td>Early dactylitis and leukocytosis</td>
<td>3 (2.5)</td>
</tr>
<tr>
<td>Severe anemia and leukocytosis</td>
<td>1 (0.5)</td>
</tr>
</tbody>
</table>

*Dactylitis before 1 year of age.
†Mean hemoglobin concentration less than or equal to 7 g/dL in second year of life.
‡Mean leukocyte count of 20,000/\text{mm}^3 or greater in second year of life.
§Combinations associated with a high probability (\(\geq 0.36\)) of severe disease by 10 years of age in the CSSCD infant cohort.

### Table 3. DNC subjects who had adverse events

<table>
<thead>
<tr>
<th>Adverse events</th>
<th>Subjects, N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Death*</td>
<td>2 (1.2)</td>
</tr>
<tr>
<td>Stroke</td>
<td>14 (8.3)</td>
</tr>
<tr>
<td>2 or more painful events per year</td>
<td>4 (2.4)</td>
</tr>
<tr>
<td>1 or more episodes of ACS per year</td>
<td>3 (1.8)</td>
</tr>
<tr>
<td>Total†</td>
<td>23 (13.7)</td>
</tr>
</tbody>
</table>

*One subject died of bacterial sepsis; the other died of complications resulting from recurrent strokes.
†Two subjects experienced two adverse events each. Only their first adverse events are included here.
the DNC. There are no publications that report a direct test of the CSSCD model in other cohorts, so a comparison of results across several cohorts cannot be made.

It is notable that we found no association between severe anemia and later adverse outcomes. This could simply be the consequence of the small number of DNC subjects who had severe anemia (N = 5). However, in accordance with the CSSCD model, we used Hgb concentration as a dichotomous predictor (Hgb < 7 g/dL or not) and considered only values from the second year of life. These restrictions certainly decreased the potential power of this variable as a predictor. Moreover, other studies have shown that steady-state Hgb concentration has contrasting features as a predictor: low values are associated with death and stroke, and high values are associated with pain and ACS. So, using a low value of Hgb concentration (Hgb < 7 g/dL) to predict all 4 adverse outcomes is likely to be problematic and yield different results in different analyses.

Our study has several limitations. First, unlike the CSSCD, DNC subjects and families were not systematically questioned about the occurrence of painful events (other than dactylitis) that were managed entirely at home or in outpatient facilities (eg, emergency rooms or clinics). We considered only painful events that required hospitalization. This could have resulted in misclassification, thereby reducing the performance of the model. For example, we treated 7 subjects with hydroxyurea who did not meet the strict CSSCD criteria for frequent pain or recurrent ACS. Perhaps the inclusion of these subjects’ outpatient-only events could have changed their classification. It is also possible that the CSSCD criteria are simply too strict, and we do not wait for this degree of disease severity to manifest before starting hydroxyurea. Second, our rate of frequent pain and recurrent ACS was lower than the CSSCD. The CSSCD used very stringent criteria that very few of our patients happened to meet, perhaps because we could not include outpatient-only episodes of pain in our analysis. Third, laboratory testing was not mandated by a protocol; instead, it was obtained according to routine clinical care and assembled retrospectively. Hematologic data were based on a median of 3 blood counts in the DNC and 4 in the CSSCD. So, some hematologic data may be missing or less representative than the data from the CSSCD. Fourth, although our clinical database prospectively records the complications of DNC subjects as part of routine clinical practice, the documentation of outcomes may not have been as rigorous and complete as that described by the CSSCD. Nevertheless, most clinicians who attempt to use this model in clinical practice will rely on the same type and quality of clinical data that we have collected and analyzed here.

Another important difference is the shorter mean follow-up time of the DNC compared with the CSSCD (7.1 vs 10 years). Because the calculation of expected probabilities incorporates follow-up time, the distribution of expected probabilities in the DNC will differ from CSSCD. Nevertheless, we found no optimal tradeoff between sensitivity and specificity for any calculated probability, and the area under the ROC curve was less than 0.5. The sensitivity of the model did not rise above 20% until the specificity fell below 60%. Perhaps an additional 2.9 years of follow-up in the DNC would have allowed time for additional adverse events to occur, thereby improving the performance of the model. However, we found that the mean time to an adverse event was 4.7 years (median, 3.7 years; interquartile range, 2.3-6.0 years). Because 75% of the adverse events occurred in first 6 years of life, an additional 2.9 years of follow-up would be unlikely to substantively change our findings. Moreover, the mean follow-up of the CSSCD validation cohort was only 5.8 years.

The main strength of this study is the use of an independently assembled cohort to test a statistical model, which is rigorous epidemiologic practice. The CSSCD model-generating and validation cohorts were subsets of a single cohort. Our sample size (N = 168), although smaller than the CSSCD model-generating cohort (N = 392), was larger than the CSSCD model-validation cohort (N = 111). Another strength is that all DNC subjects were identified by universal newborn screening for hemoglobinopathies, and we prescribed prophylactic penicillin to all children with SS or Sβ+ until at least 5 years of age. Use of newborn screening and penicillin was not universal in the CSSCD infant cohort. The consequently higher frequency of fatal bacterial sepsis in the CSSCD cohort (10 of 18 deaths) is discordant with current patterns of mortality in SCD. Perhaps the model performed poorly here, in part, because our contemporary cohort has a lower incidence of infection-related death (1 of 2 deaths). Stroke was the most frequent adverse event in both the DNC and CSSCD, occurring in 8.3% and 6.3% of subjects, respectively. Our data from the DNC, like the CSSCD, do not reflect the relatively recent impact of transcranial Doppler screening, so the frequency of clinically overt stroke will continue to decrease with time. It is possible that the CSSCD model does predict adverse outcomes, especially death and stroke, but not the adverse outcomes that are relatively more common today. As such, the model may no longer be prognostic in the setting of current management. A better measure of disease severity, beyond the occurrence of death and stroke, is clearly needed.

In summary, we have failed to validate the CSSCD early prediction model in an independent cohort of young children
with SCD. The model could not identify the children who would later have adverse outcomes based on a combination of early clinical features. It should not be used as the sole criterion to initiate early, high-risk interventions. A robust early prediction model is still needed. However, there are at least 2 barriers to its development. First, we need a meaningful measure of disease severity. Adverse outcomes like death and stroke are increasingly rare in children, so we now require a contemporary definition of the severity of SCD. Second, it may not be possible to use simple clinical and laboratory markers as predictors of such a complex disease. Broad genomic and proteomic approaches should be explored to advance this field.

Acknowledgments

We thank Scott T. Miller, MD, and Lynn A. Sleeper, ScD, from the CSSCD for providing the parameter estimates for the predictive model and for their review of the manuscript.

Authorship

Conflict-of-interest disclosure: The authors declare no competing financial interests.

Correspondence: Charles T. Quinn, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75390-9063; e-mail: charles.quinn@utsouthwestern.edu.

References


This work was supported in part by grants L40 HL074825 and U54 HL70588 from the National Institutes of Health.
Prediction of adverse outcomes in children with sickle cell anemia: a study of the Dallas Newborn Cohort

Charles T. Quinn, Nancy J. Lee, Elizabeth P. Shull, Naveed Ahmad, Zora R. Rogers and George R. Buchanan