Inter-Hospital Variation in Growth Outcomes and Improvement Using a Quantitative Metric

Kristyn Beam MD, Sergei Roumiantsev MD PhD, Terri Gorman MD, Munish Gupta MD, Nneka Nzegwu DO, Sunita Pereira MD, Alan Picarillo MD, Rachana Singh MD, Andrew L Beam PhD, Joseph H Chou MD PhD

Primary author: Joseph H Chou, Massachusetts General Hospital, jchou2@partners.org 617-724-9040

Keywords: growth, nutrition, quality improvement, quantitative

Background: Prenatal growth failure is common in premature infants and may be associated with long-term neurodevelopmental impairment. We hypothesized that a simple metric to assess growth outcomes—the change in weight Z score from birth to discharge (ΔZ)—might yield insight into growth outcome variation between NICUs and provide a metric to assess the efficacy of nutrition quality improvement interventions.

Aim: 1) Demonstrate the feasibility of change in weight Z-score to assess growth outcomes, and to compare outcomes between hospitals; 2) At NICU C, implement bundled growth improvement interventions to reduce the loss in weight Z-score from birth to discharge, demonstrating sustained improvement starting in year one.

Setting: All infants born between 23 0/7 and 34 6/7 weeks gestational age from 2006 – 2011 at seven level three NICUs in Massachusetts with Vermont Oxford Network (VON) registry data available were used for growth outcomes analysis. Post-intervention data through 2014 was included for NICU C for Aim 2.

Mechanism: Factors potentially contributing to impaired growth were identified (Figure 1).

Drivers of Change and Methods: Bundled interventions targeting factors contributing to growth failure were implemented at NICU C in 2011, and included raised awareness of baseline growth failure; development of an electronic growth chart; weekly growth metric collection and review; earlier and broader initiation of parenteral nutrition; increased protein content and more rapid advancement of parenteral nutrition; revision of enteral feeding advancement protocols; and development of a web calculator for growth metrics (http://pedtools.org/).

Measures: Patient level data were abstracted from VON datasets at each hospital, including initial gestational age, birth weight, length of stay, and discharge weight. The outcome metric was change in the weight Z-score (ΔZ) from birth to discharge, and variation between hospitals and by gestational age were analyzed.

Data and Results: 7,975 neonates born between 23 and 34 weeks gestation were analyzed. Patient populations at the 7 NICUs differed significantly (Table 1). Growth failure was demonstrated with a mean decrease in weight Z-score of 0.81 from birth to discharge which was inversely related to gestational age, with each additional week of prematurity further lowering the ΔZ score by 0.029 (p < 0.0001; Figure 2).

1) Inter-hospital comparison of outcomes: there were differences in growth outcomes between hospitals, with ΔZ weight ranging from −1.05 to −0.56 (p < 0.0001; Table 1). The relationship between gestation at birth and ΔZ also differed by hospital (Figure 3). Patterns of growth at individual hospitals remained stable across different birth year epochs (data not shown) suggesting reliability for use as a quality improvement metric.

2) ΔZ weight as a quality improvement outcome metric: at NICU C, bundled interventions introduced in 2011 targeting improved growth and nutrition resulted in significant improvement, particularly at the lowest gestational ages, and this improvement was sustained and progressive (Figure 4).

Discussion: Growth outcomes during the birth hospitalization of a large population of premature newborns from 23 to 34 weeks gestational age at seven Massachusetts NICU’s were analyzed by change in weight Z score from birth to discharge, demonstrating significant growth failure during the birth hospitalization; worsened growth failure with increasing prematurity; and significant intersite differences in degree of growth failure. This metric was effective as an outcome measure to demonstrate improvement in growth outcomes after process improvement interventions, suggesting potential use as a quantitative metric to benchmark individual NICU growth outcomes as well as to facilitate comparison between different NICUs.
Table 1. Study population of 7,975 premature newborns born between 23 and 34 weeks completed weeks gestation, in 7 hospitals (A – G). Weight delta Z is the change in Z-score for weight from birth to discharge (Fenton 2003). An additional 570 neonates born between 2012 and 2014 from NICU C were included for post-intervention outcomes analysis, not tabulated here.

![Figure 1. Potential contributors to impaired growth outcomes in premature newborns. An intervention bundle targeting these causes included: raised awareness of baseline growth failure; weekly growth metric collection and review; development of an electronic growth chart; earlier and broader initiation of parenteral nutrition; increased protein content and more rapid advancement of parenteral nutrition; availability of pasteurized donor human milk; revision of enteral feeding advancement protocols; and development of a web tool to calculate growth metrics (http://peditools.org/)](http://peditools.org/)

### Table 1

<table>
<thead>
<tr>
<th></th>
<th>All NICUs</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of patients</td>
<td>7,975</td>
<td>461</td>
<td>1,586</td>
<td>1,068</td>
<td>418</td>
<td>598</td>
<td>1,081</td>
<td>2,763</td>
<td>–</td>
</tr>
<tr>
<td>Gestational age (median)</td>
<td>32</td>
<td>29.0</td>
<td>33.0</td>
<td>32.9</td>
<td>31.9</td>
<td>29.1</td>
<td>31.0</td>
<td>32.6</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Birth weight (g, median)</td>
<td>1,580</td>
<td>1,130</td>
<td>1,810</td>
<td>1,819</td>
<td>1,463</td>
<td>1,160</td>
<td>1,400</td>
<td>1,750</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Birth weight Z-score (mean)</td>
<td>–0.19</td>
<td>–0.40</td>
<td>–0.14</td>
<td>–0.07</td>
<td>–0.32</td>
<td>–0.36</td>
<td>–0.24</td>
<td>–0.14</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Weight delta Z (mean)</td>
<td>–0.81</td>
<td>–0.75</td>
<td>–1.05</td>
<td>–0.80</td>
<td>–0.58</td>
<td>–0.96</td>
<td>–0.56</td>
<td>–0.79</td>
<td>&lt; 0.0001</td>
</tr>
</tbody>
</table>
Figure 2. Plot of change in weight $Z$-score from birth to discharge versus gestational age at birth, demonstrating worsening growth failure with increasing prematurity. The solid line shows a scatterplot smoother regression fitting with the shaded area demonstrating the 95% confidence interval of the estimated mean. The superimposed dotted line show the same data fitted by linear regression.
Figure 3. Plot of change in weight Z-score from birth to discharge versus gestational age at birth, by each of seven different hospital NICUs in Massachusetts. There are significant inter-hospital differences in both the degree of growth failure as well as the dependence on gestational age at birth. Some hospitals show much higher growth failure at lower gestational ages at birth (e.g., hospital E), whereas other hospitals show less growth failure overall and absence of worse outcomes at lower gestational age (e.g., hospital F). The lines show a scatterplot smoother regression fitting with the shaded area demonstrating the 95% confidence interval of the estimated mean.
Figure 4. Improvement of growth outcomes after implementation of an intervention bundle targeting improved nutrition and growth. **A.** Mean change in weight Z-score by year at hospital C of babies born before 32 weeks gestation, via Xbar statistical process control chart and 3-SD control limits. The vertical line represents when the interventions were implemented. **B.** Plot of change in weight Z-score from birth to discharge versus gestational age at birth at Hospital C, by birth year epoch. 2006 – 2010 represents a pre-intervention baseline epoch; 2011 – 2012 is the first post-intervention epoch; 2013 – 2014 demonstrates sustained improvement. The largest improvements are seen at the lowest gestational ages at birth. The lines show a scatterplot smoother regression fitting with the shaded area demonstrating the 95% confidence interval of the estimated mean.