ARTICLE

Quality improvement project designed to reduce disproportionate growth in extremely low gestational age neonates: cognitive neurodevelopmental outcome at 18–41 months

Jordan D. Reis[1] · Kristine Tolentino-Plata[1] · Roy Heyne[1] · L. Steven Brown[2] · Charles R. Rosenfeld [1] ·
Maria Caraig[1] · Patti J. Burchfield[1] · Luc P. Brion [1]

Abstract

Objective To assess if the adjusted odds of low composite cognitive Bayley-III scores changed after implementing a single-institution quality improvement (QI) project designed to decrease discharge Z-scores for weight, body mass index (BMI), and weight-for-length, but not length or fronto-occipital circumference (FOC) in infants 23–28 weeks gestational age (GA).

Methods Compare Bayley-III outcomes at ≥18 months corrected age (postnatal age adjusted for prematurity) in infants tested before (Epoch-1) and after (Epoch-2) QI implementation.

Results Bayley testing was available in 134/156 infants (86%) in Epoch-1 and 139/175 (79%) in Epoch-2. There was no change in frequency of low (<85) cognitive score ($p = 0.5$) or in median cognitive scores (80 in Epoch-1 vs. 85 Epoch-2, $p = 0.35$). The adjusted odds of low cognitive scores was not different between Epochs.

Conclusion No change in cognitive outcome at ≥18 months corrected age was observed after implementing a QI project designed to reduce discharge weight-for-length disproportion in very preterm infants.

Introduction

More than one in ten infants in the United States are born preterm each year. Those born very preterm (≤28 weeks gestational age (GA)) have the highest risk of neonatal mortality and morbidity and unfortunately, the greatest burden of neurodevelopmental disabilities. Much research has therefore been devoted to identifying ways to improve long-term neurodevelopmental outcomes in this vulnerable population.

Greater postnatal growth in preterm infants has long been associated with improved neurodevelopmental outcomes [1–9]. However, this has generally focused on increasing weight gain velocity. This has been due to a lack of consensus on a safe, valid method for accurately measuring serial lengths in critically ill neonates. This approach to judging growth can lead to disproportionate weight-for-length growth with excessive weight gain and body mass index (BMI) at the time of discharge from the neonatal intensive care unit (NICU) and later in infancy, potentially contributing to the increased frequency of metabolic syndrome in young adults born preterm [10–13]. Recent studies have shown a stronger relationship between neurodevelopmental outcome and head growth and linear growth velocities than weight gain velocity [14–19].

The aim of this study was to compare neurodevelopmental outcome of very preterm infants before (Epoch-1) and after (Epoch-2) implementing a quality improvement (QI) project designed to reduce disproportionate growth (weight > length) from birth to discharge from the NICU [20, 21]. Infants born in Epoch-2 had lower discharge Z-scores for weight, weight-for-length, and BMI compared with those born in Epoch-1, but similar Z-scores for length and fronto-occipital circumference (FOC) [20]. We hypothesized that cognitive outcomes in

✉ Luc P. Brion
luc.brion@utsouthwestern.edu

1   Division of Neonatal-Perinatal Medicine, Department of Pediatrics, University of Texas Southwestern Medical Center, Dallas, TX, USA

2   Parkland Health and Hospital System, Dallas, TX, USA

very preterm infants born in Epoch-2 would not be worse than those born in Epoch-1.

Methods

Study design

Previous studies of preterm very low birth weight (VLBW) infants born at Parkland Hospital had shown disproportionate weight-for-length (>90th centile) growth, greater waist circumference and fat mass at NICU discharge and persistence at 1 and 2 years of age [10–12]. At 2-year follow-up, these infants also had elevated serum adipokines, while at 10–13 years (studies in analysis) they had a high frequency of excessive weight gain and adiposity associated with an elevated systolic blood pressure [13].

In order to limit the occurrence of disproportionate weight-for-length growth at NICU discharge, we initiated a QI project targeting neonates born at <29 weeks GA and small for GA neonates born at <34 weeks GA at Parkland Hospital that included implementation of a validated method for serial length measurements and adjustable nutrient intake based on serum laboratory measurements [20, 21]. Neonates were individually fortified based on growth and serial serum concentrations of key nutrients and five growth patterns based on analysis of weight gain and linear growth: physiologic growth, insufficient growth, insufficient linear growth, excessive weight gain, and insufficient weight gain. Infants with insufficient linear growth, which was associated with deficiency in at least one nutrient (e.g., protein, electrolytes, minerals, vitamins, and/or zinc), received individual fortification to address such deficiency in order to improve linear growth and head growth [21, 22]. Implementation of the QI project resulted in decreased discharge Z-scores for weight, BMI, and weight-for-length, but not discharge length or FOC Z-scores in appropriate for GA neonates <29 weeks GA. Growth and head growth after discharge were associated with human milk (HM) at discharge and time of initiation and type of complementary food intake [23]. The current study reports on the Bayley-III neurodevelopmental testing results obtained on neonates <29 weeks GA when they were seen in our NICU follow-up clinic at ≥18 months adjusted age (postnatal age corrected for prematurity).

Setting

All infants were born at Parkland Hospital, a large public hospital with >12,000 annual births. They were followed in the Thrive Clinic at Children's Medical Center of Dallas. The Thrive Clinic provides comprehensive primary care to high-risk infants in addition to developmental and focused research assessments. Developmental assessment included the Bayley Scales of Infant Development III performed by a certified examiner (KTP) in either English or Spanish [24].

Inclusion and exclusion criteria

We included infants $23^{0/7}$–$28^{6/7}$ weeks GA born between July 2012 and March 2015 (Epoch-1) or between June 2015 and October 2017 (Epoch-2) who were discharged from our NICU and then completed Bayley-III neurodevelopmental testing in the Thrive Clinic at ≥18 months adjusted age. Infants who died after NICU discharge were lost to follow-up, or who did not complete Bayley-III neurodevelopmental testing at ≥18 months adjusted age were excluded.

Variables

The primary outcome variable was the adjusted odds of low composite cognitive score (<85). Variables considered for the multivariate analyses included: (1) Epoch-2 vs. Epoch-1; (2) maternal data: education level, race/ethnicity, multiple pregnancy, antenatal steroids, (3) neonatal data at birth: GA, birth weight Z-score calculated using Olsen's growth curves [25], sex; (4) neonatal morbidity: moderate or severe bronchopulmonary dysplasia (BPD) [26], severe intraventricular hemorrhage (IVH) ≥grade III [27], periventricular leukomalacia, necrotizing enterocolitis (NEC) stage ≥II [28], severe retinopathy of prematurity (ROP) (defined as retinal detachment, laser surgery, or exposure to local injection of Bevacizumab, an anti-vascular endothelial growth factor); (5) discharge data: use of HM, Z-scores of weight, length, FOC, and BMI calculated using Olsen's or WHO curves [25, 29, 30], changes in weight, length and FOC Z-scores from birth to discharge; (6) language (Spanish, English or other); and (7) adjusted age at neurodevelopmental testing.

Secondary outcome variables included low composite language score (<85), low composite motor score (<85), very low composite motor score (<70), as well as cognitive, language, and motor scores assessed as continuous variables.

Statistical analysis

Analysis was performed using SPSS version 26 (IBM, Inc., Armonk, NY). Demographic and other clinically relevant information are summarized with percentages, means and standard deviations (SD), or medians and interquartile ranges (IQR). Continuous data were compared using Student t test, Welch test (used if Levene's test showed unequal variances), or Mann–Whitney U test depending on the data distribution. Categorical data were compared with a Chi-square (χ^2) test or Fisher exact test. Composite Bayley

Table 1 Birth characteristics of infants included in Epochs-1 and -2.

Birth characteristics	Epoch-1 (n = 134)	Epoch-2 (n = 139)	p value[c]
Maternal education level			
<High school education	61 (45.5)[a]	58 (41.7)	0.30
High school graduate	22 (16.2)	33 (23.7)	
Partial college or trade	30 (22.4)	25 (18)	
≥College graduate	10 (7.5)	16 (11.5)	
Unknown	11 (8.2)	7 (5)	
Antenatal steroids	74 (55)	118 (85)	<0.001
Gestational age (weeks)	26.2 ± 1.3[b]	26.2 ± 1.4	0.69
Female	65 (49)	73 (53)	0.51
Singleton	107 (80)	115 (83)	0.54
Weight (g)	907 ± 213	907 ± 245	0.99
Length (cm)	34 ± 3	34 ± 3	0.45
Fronto-occipital circumference (cm)	24 ± 2	24 ± 2	0.79
Weight Z-score	0.11 ± 1.03	0.19 ± 1.28	0.59
Length Z-score	0.00 ± 1.04	−0.05 ± 1.18	0.71
FOC Z-score	−0.13 ± 1.10	−0.10 ± 1.36	0.86
Race/ethnicity			0.03[d]
Hispanic White	95 (71)	89 (64)	
Non-Hispanic African–American	29 (22)	45 (32)	
Non-Hispanic White	5 (4)	5 (4)	
Other	5 (4)	0 (0)	
Language			
Spanish	65 (49)	55 (40)	0.11[d]
English	65 (49)	83 (60)	
Other	4 (3)	1 (1)	

FOC fronto-occipital circumference.

[a]Values are number and percent of the total in each column.

[b]Values are mean ± SD.

[c]Student *t* test or χ^2 analysis.

[d]Exact test.

scores have a mean of 100 and a SD of 15 [31]. The definition of low Bayley scores is arbitrary and controversial and depends on each specific version [32]. Therefore, we analyzed data as dichotomous (logistic regression) and continuous (linear regression). Multivariate logistic regression with forward stepwise inclusion (in, $p < 0.05$, out, $p < 0.10$) was used to model low and very low composite Bayley scores and to calculate the adjusted odds ratio and 95% confidence intervals of low Bayley scores. Multivariate linear regression with backward stepwise elimination was used to model Bayley scores with covariates. In addition, the number of independent variables for each model was limited to one variable per ten patients with the outcome of interest and to eliminate multi-collinearity, which was assessed using SAS 9.4 (SAS Institute, Inc,

Cary, NC) [33, 34]. All statistical tests used <0.05 as the significance level.

The sample size of this study was larger than that of the original study ($n = 226$) [21], which excluded small for GA and large for GA infants; however, it was limited by the number of infants who had complete Bayley scores. The current study had 80% power calculated using SAS with a two-tailed alpha of 0.05 to detect an association of Epoch with a 33% increase in percentage of low cognitive Bayley scores (66.7% in Epoch-2 vs. 50.0% in Epoch-1, corresponding to an odds ratio of 2 for the primary outcome) using logistic regression analysis adjusted for one continuous and two dichotomous independent variables. Secondary analyses should be considered exploratory because we had insufficient power and did not adjust p values for multiple comparisons.

Results

Study population

The study population included 331 infants <29 weeks GA discharged from the NICU (Supplementary Fig. 1). Bayley testing at 18–41 months adjusted age was available in 134/156 infants (86%; 26.2 ± 1.3 weeks GA) in Epoch-1 and 139/175 (79%; 26.2 ± 1.4 weeks GA) in Epoch-2 (Table 1).

The most common reasons for exclusion included loss to follow-up after NICU discharge ($n = 34$) and performance of neurodevelopmental testing before 18 months adjusted age ($n = 15$) (Supplementary Fig. 1). In both Epochs, infants who completed Bayley testing at ≥18 months adjusted age had lower mean GA and Z-scores for birth weight, length, and FOC than those who did not have complete Bayley testing (Supplementary Table 1). In Epoch-1, infants who completed Bayley testing at ≥18 months had lower mean Z-scores for length and FOC at discharge compared to those in Epoch-1 who did not complete Bayley testing. In Epoch-2, infants who completed Bayley testing at ≥18 months were discharged at later postmenstrual age with lower Z-scores for weight and length compared to those in Epoch-2 who did not complete Bayley testing (Supplementary Table 2).

The race/ethnic distribution was different and the percentage of administration of antenatal steroids was higher in Epoch-2 compared with Epoch-1 (Table 1). The frequencies of severe NICU morbidities were similar for the two Epochs except for a higher percentage of infants with severe ROP in Epoch-1 (Table 2a). A higher percentage of infants in Epoch-1 received HM feeds at time at the of NICU discharge compared with those in Epoch-2 (48% vs. 36%, respectively, $p = 0.048$; Table 2b); this was likely related to changes that occurred after the move of the NICU to a

Table 2 NICU morbidities and discharge characteristics in infants included in Epochs-1 and -2.

	Epoch-1 (n = 134)	Epoch-2 (n = 139)	p value[d]
(a) NICU morbidities			
PDA ligation	15 (11)[b]	10 (7)	0.25
BPD (moderate or severe)	40 (30)	36 (26)	0.47
NEC (Stage II–III)	11 (8)	8 (6)	0.43
IVH (Grade III–IV)	18 (13)	14 (10)	0.39
Cystic PVL	6 (5)	2 (1)	0.17
Severe ROP[a]	27 (20)	15 (11)	0.03
(b) Discharge characteristics			
Human milk feeds (yes)	64 (48)	50 (36)	0.048
Weight (g)	3444 ± 684[c]	3248 ± 854	0.04
Length (cm)	49 ± 3	49 ± 4	0.97
FOC (cm)	35 ± 2	34 ± 2	0.62
Weight Z-score	−0.57 ± 0.89	−1.01 ± 0.88	<0.001
Length Z-score	−1.53 ± 1.25	−1.54 ± 1.19	0.96
FOC Z-score	−0.66 ± 1.07	−0.77 ± 1.16	0.39
BMI	14.52 ± 1.52	13.57 ± 1.76	<0.001
BMI Z-score	0.53 ± 0.81	−0.11 ± 0.84	<0.001
Weight-for-length Z-score[e]	1.06 ± 1.06 (n = 124)	0.27 ± 1.22 (n = 131)	0.001
Change in weight Z-score from birth to discharge	−0.68 ± 0.82	−1.19 ± 1.07	<0.001
Change in length Z-score from birth to discharge	−1.51 ± 1.00	−1.49 ± 1.07	0.93
Change in FOC Z-score from birth to discharge	−0.54 ± 1.00	−0.69 ± 1.31	0.31
Postmenstrual age (weeks)	40.7 ± 3.9	40.9 ± 4.7	0.66

PDA patent ductus arteriosus, *BPD* bronchopulmonary dysplasia, *NEC* necrotizing enterocolitis, *IVH* intraventricular hemorrhage, *PVL* periventricular leukomalacia, *ROP* retinopathy of prematurity, *FOC* front-occipital circumference, *BMI* body mass index.

[a]Stage III–IV and/or requiring Bevacizumab or laser therapy.

[b]Values are number and percent of the total in each column.

[c]Values are mean ± SD.

[d]Student *t* test, χ^2 analysis or Fisher's exact test.

[e]Limited to postmenstrual age ≥37 weeks and length ≥44 cm.

newly built hospital (change in breast pump supplier, much larger unit with individual patient rooms located on a separate floor from labor and delivery and postpartum units, insufficient number of refrigerators available for breast milk storage within the postpartum unit, limited number of personnel available for breast milk transport).

As expected, given the aim of the QI project, infants in Epoch-1 had a significantly higher weight and weight Z-score at NICU discharge compared to infants in Epoch-2 (3444 ± 684 g vs. 3248 ± 854 g, $p = 0.04$, and −0.57 ± 0.89 vs. −1.01 ± 0.88, respectively; $p < 0.001$) (Table 2b). In contrast, their length and FOC measurements and Z-scores did not differ. In lieu of this, BMI and BMI Z-scores and weight-for-length Z-scores at time of NICU discharge were significantly lower in Epoch-2 (Table 2b). The decrease in weight Z-scores from birth to discharge was greater in Epoch-2 than in Epoch-1, whereas changes in length Z-scores and FOC Z-scores were similar in both Epochs (Table 2b).

Neurodevelopment

Bivariate analyses

Neurodevelopmental outcomes assessed by Bayley-III scores are shown in Table 3. Testing was performed at a similar adjusted age in both Epochs. There was no difference in frequency of low composite cognitive score ($p = 0.51$), low composite motor scores ($p = 0.21$), or very low motor scores ($p = 0.23$) in Epoch-2; however, there was a significant increase in the frequency of low composite language score in Epoch-2 ($p = 0.02$; Table 3).

Notably, Bayley scores were skewed to the left in both Epochs and did not fit a Gaussian distribution; therefore, the main comparison was done with nonparametric statistics. There was no significant change in median composite cognitive score between the two Epochs ($p = 0.35$). There was a significant decrease in median composite language scores [79 (IQR 71, 89) in Epoch-1 vs. 74 (IQR 65, 84) in Epoch-2, $p = 0.003$] and increase in median composite motor scores in Epoch-2 [88 (IQR 79, 94) vs. 91 (IQR 82, 97), $p = 0.006$, respectively].

Logistic regression analyses

The odds of low composite cognitive scores was negatively associated with the use of HM at discharge ($p = 0.01$) and positively associated with severe IVH ($p = 0.01$) and moderate or severe BPD ($p = 0.009$); it was unchanged with Epoch (Table 4a).

The odds of low composite language scores was associated with an interaction between Epoch and race/ethnicity ($p = 0.009$), negatively associated with age at testing ($p = 0.001$) and positively associated with male sex ($p = 0.001$) and severe ROP ($p = 0.009$) (Table 4b).

The odds of low composite motor scores was negatively associated with use of HM ($p = 0.008$) and Z-score of FOC at discharge ($p = 0.004$) and positively associated with severe IVH ($p = 0.005$) and moderate or severe BPD ($p = 0.01$); it was unchanged with Epoch (Table 4c).

Table 3 Bayley scores obtained at ≥18 months follow-up in infants included in Epochs-1 and -2.

Bayley scores	Epoch-1 ($n = 134$)	Epoch-2 ($n = 139$)	p value[c]
Adjusted age (months)	22.4 (22.1, 23.7)[a]	22.4 (22.1, 23.3)	0.34
Cognitive score: median (IQR)	80 (75, 90)	85 (75, 90)	0.35
Language score: median (IQR)	79 (71, 89)	74 (65, 84)	0.003
Motor score: median (IQR)	88 (79, 94)	91 (82, 97)	0.006
Cognitive score <85	70 (52)[b]	67 (48)	0.51
Language score <85	83 (62)	104 (75)	0.02
Motor score <85 (low)	47 (36)	40 (29)	0.21
Motor score <70 (very low)	15 (12)	10 (5)	0.23

IQR interquartile range.

[a]Values are median (interquartile range).

[b]Values are number and percent of the total in each column.

[c]Mann–Whitney test, χ^2 analysis or Fisher's exact test.

Table 4 Adjusted odds of low cognitive and language Bayley scores occurring at 18–41 months follow-up in infants included in Epochs-1 and -2.

Factor	Adjusted odds ratio	95% confidence interval	p value
(a) Cognitive score <85 ($n = 135$ among 269)[a]			
Human milk at discharge	0.54	0.32, 0.88	0.01
Severe IVH	3.04	1.29, 7.18	0.01
Moderate or severe BPD	2.11	1.20, 3.70	0.009
(b) Language <85 ($n = 187$ among 273)			
Male vs. female	2.60	1.46, 4.63	0.001
Interaction between Epoch and race/ethnicity	–	–	0.003
Severe ROP	3.88	1.40, 10.78	0.009
Adjusted age at testing (by month)	0.89	0.82, 0.95	0.001
(c) Motor score <85 ($n = 86$ among 272)			
Human milk at discharge	0.45	0.25, 0.81	0.008
Severe IVH	2.79	1.25, 6.20	0.005
Moderate or severe BPD	2.14	1.17, 3.90	0.01
Z-score of FOC at discharge	0.69	0.53, 0.89	0.004
(d) Motor score <70 ($n = 25$ among 273)			
Severe IVH	8.36	3.25, 21.45	<0.001
Severe ROP	2.93	1.11, 7.75	0.03

Other variables did not reach significance.

BPD bronchopulmonary dysplasia, *IVH* intraventricular hemorrhage, *ROP* retinopathy of prematurity.

[a]Area under the curve: 0.66. Stepwise forward logistic regression analysis ($p < 0.05$ in, $p < 0.10$ out) adjusted to yield a final model with non-significant Hosmer–Lemeshow test, non-significant collinearity diagnostics and maximum 1 variable per 10 infants with the outcome of interest [low Bayley score]). Variables that were considered for inclusion in each model included Epoch, maternal education level, race/ethnicity, antenatal steroids, multiple birth, gestational age, birthweight Z-score, sex, moderate/severe BPD, severe IVH, periventricular leukomalacia, necrotizing enterocolitis, severe ROP, use of human milk at discharge, Z-scores of weight, length, BMI, and FOC at discharge, changes in weight, length, and FOC Z-scores from birth to discharge, language (Spanish, English or other) and adjusted age at neurodevelopmental testing. Variables with significant collinearity diagnostics included Z-scores of weight, length, and BMI at discharge and change in Z-scores of length and FOC from birth to discharge.

The odds of very low composite motor scores was positively associated with severe IVH ($p < 0.001$) and severe ROP ($p = 0.03$); it was unchanged with Epoch (Table 4d).

Multivariate linear regression

Cognitive scores were negatively associated with severe IVH ($p = 0.001$) and moderate or severe BPD ($p = 0.03$) and positively associated with the use of HM at discharge ($p = 0.007$); they were unchanged with Epoch (Table 5a).

Language scores were associated with an interaction between Epoch and race/ethnicity ($p < 0.001$), negatively associated with male sex ($p < 0.001$) and severe ROP ($p = 0.001$) and positively associated with use of HM at discharge ($p < 0.001$) and age at testing ($p < 0.001$); they were unchanged with Epoch (Table 5b).

Motor scores were negatively associated with severe IVH ($p < 0.001$) and positively associated with Epoch-2 ($p = 0.002$), use of HM at discharge ($p = 0.001$) and FOC Z-score at discharge ($p < 0.001$) (Table 5c).

Table 5 Multivariate linear regression of Bayley scores occurring at 18–41 months follow-up in infants included in Epochs-1 and -2.

Factor	Beta coefficient per unit of Bayley score	95% confidence interval	*p* value
(a) Cognitive score ($n = 273$)			
Human milk at discharge	3.61	1.00, 6.21	0.007
Severe IVH	−7.63	−12.06, −3.20	0.001
Moderate or severe BPD	−3.44	−6.48, −0.40	0.03
(b) Language score ($n = 271$)			
Male vs. female	−5.95	−8.89, −3.00	<0.001
Interaction between Epoch and race/ethnicity	–	–	<0.001
Human milk at discharge	6.10	3.13, 9.07	<0.001
Severe ROP	−6.82	−10.91, −2.73	0.001
Adjusted age at testing (by month)	7.26	0.33, 1.12	<0.001
(c) Motor score ($n = 269$)			
Epoch-2 vs. Epoch-1	4.43	1.57, 7.29	0.002
Human milk at discharge	4.68	1.84, 7.52	0.001
Severe IVH	−11.20	−17.38, −5.01	<0.001
FOC Z-score at discharge	2.89	1.51, 4.26	<0.001

Generalized linear model including variables that reached significance in a backward stepwise analysis limited to maximum 1 variable per 10 infants and to non-significant collinearity diagnostics. Variables that were considered for inclusion in each model included Epoch, maternal education level, race/ethnicity, antenatal steroids, multiple birth, gestational age, birthweight Z-score, sex, moderate/severe BPD, severe IVH, periventricular leukomalacia, necrotizing enterocolitis, severe ROP, human milk at discharge, Z-scores of weight, length, BMI, and FOC at discharge, changes in weight, length, and FOC Z-scores from birth to discharge., language (Spanish, English or other) and adjusted age at testing. Variables with significant collinearity diagnostics included Z-scores of weight, length, and BMI at discharge and change in Z-scores of length and FOC from birth to discharge.

BPD bronchopulmonary dysplasia, *IVH* intraventricular hemorrhage, *ROP* retinopathy of prematurity.

Discussion

Previous studies of preterm VLBW infants born at Parkland Hospital had shown disproportionate weight-for-length growth in infancy and high frequency of obesity and high systolic blood pressure at 10–13 years of age. Implementation of a QI project that reduced weight-for-length disproportion at NICU discharge among preterm infants <29 weeks GA was not associated with any significant change in the frequency of low cognitive Bayley scores or in median cognitive scores. The primary outcome, i.e., the adjusted odds of low composite cognitive score, was positively associated with severe IVH and moderate or severe BPD and negatively associated with the use HM at discharge, but unchanged with Epoch. Several secondary analyses should be considered as exploratory because of insufficient power and lack of adjustment of *p* values for multiple comparisons; these include lower language scores and higher motor scores in Epoch-2 than in Epoch-1. Thus, these data suggest there is no evidence of neurodevelopmental benefit to excessive weight gain that results in disproportionate growth in very preterm infants during their NICU hospitalization.

In our study, use of HM at discharge was associated with improved cognitive, language, and motor Bayley scores. HM is considered the gold standard source of nutrition for preterm infants, and there are both short- and long-term benefits known to be associated with its use, including decreased rates of NEC, late-onset sepsis, and ROP, fewer re-hospitalizations in the first year of life, and improved neurodevelopmental outcomes [35, 36]. Notably, the "breastfeeding paradox" was identified in an observational study of ~3000 infants, which demonstrated that very preterm infants who were breastfeeding at the time of NICU discharge had, in comparison with those who were not breastfeeding at discharge, an increased risk of a having a decrease in weight Z-score by more than one unit from birth to discharge [16]. However, breastfeeding at discharge was associated with a decreased risk for suboptimal neurodevelopmental assessment at 2 or 5 years adjusted age and an increased risk of having a higher FOC Z-score at follow-up [16]. Length and weight-for-length were not provided in that study. Patra et al. observed a dose-dependent positive association between in-hospital HM intake in VLBW infants and their cognitive outcomes at 20 months corrected age [37]. Although a significantly greater number of infants were discharged home with HM feedings in Epoch-1 in our study, the cognitive Bayley scores did not differ between the two Epochs.

In our study there also was no association between cognitive Bayley scores and the change in weight Z-score from birth to discharge from the NICU; this is in agreement with a recent review of the literature [38]. An association of worse neurodevelopment with disproportionate growth defined by high BMI Z-scores at discharge was reported in a

large observational study in extremely preterm infants with poor linear growth [14]. In addition, we observed an association between increasing the Z-score of FOC at discharge and improved motor Bayley scores. Similarly, the association of neurodevelopment with FOC growth in the NICU and after discharge has been reported in several studies [16–19]. An association between neurodevelopment and linear growth has also been reported by several investigators [14, 15, 39, 40]. In summary, our observations are in agreement with the literature that there is an association of neurodevelopment with head growth rather than weight gain in the NICU.

Strengths of this QI project include the fact that it involved a nutritional intervention and was not simply an observational study. Multivariate analyses allowed us to study changes in neurodevelopment in association with changes in weight gain independently of severity of illness. Lack of change in cognitive scores with Epoch was shown both in unadjusted analyses and adjusted dichotomous and continuous analyses. Moreover, it included accurate serial length measurements in the prospective arm (Epoch-2), thereby allowing us to identify infants with abnormal growth patterns in weight, length, and weight-for-length proportion. This, in turn, allowed us to measure specific serum nutrients in infants who demonstrated abnormal growth in three parameters and thus, to identify deficiencies that are malleable to appropriate protein and micronutrient supplementation. Emerging evidence suggests that very preterm infants are at increased risk for micronutrient deficiency, including zinc, and that zinc supplementation can improve FOC growth, brain growth, and neurodevelopment [23, 41].

Since this study had a quasi-experimental design, associations rather than cause-and-effect relationships were analyzed, with multivariate analyses to adjust for covariates. Infants who had Bayley-III scores had lower GA and were smaller at birth and discharge than those who were not tested. This difference might have caused a selection bias, thereby impacting the frequency of low Bayley scores in this study. Further limitations include the partially retrospective component of the project and lack of data on body composition of infants at time of NICU discharge and follow-up. Lack of accurate birth lengths in most infants in Epoch-1 limited our ability to accurately compare changes in length Z-scores from birth to discharge between the two Epochs. Several secondary analyses should be considered as exploratory because of insufficient power and lack of adjustment of p values for multiple comparisons. The sample size was too small to determine whether there was an interaction between changes in language scores, race/ethnicity, primary language spoken at home and bilingualism; this effect was tentatively minimized by the fact that the Bayley examiner was fluent in English and Spanish, which were the most frequent languages.

In summary, we describe Bayley-III scores in infants born $23^{0/7}$–$28^{6/7}$ weeks GA who participated in a QI project designed to decrease weight-for-length disproportion at NICU discharge without affecting discharge length and head circumference Z-scores. Infants born in the Epoch-2 showed, compared to those born in Epoch-1, clinically similar cognitive, but lower language and higher motor neurodevelopmental outcomes at 18–41 months adjusted age on bivariate analyses. Thus, there is no evidence of neurodevelopmental benefit to excessive weight gain that results in disproportionate growth in very preterm infants during their NICU hospitalization. At present, we continue to follow these infants in order to evaluate the expression of various components of metabolic syndrome since we originally hypothesized that infants with decreased BMI and weight-for-length disproportion at NICU discharge might have a decreased incidence of obesity, hypertension, and elevated serum adipokines at long-term follow-up.

Data availability

The datasets generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

Compliance with ethical standards

Conflict of interest The authors declare no competing interests.

Ethical approval The original QI project, including follow-up, received approval by the Institutional Review Board (IRB) of the University of Texas Southwestern Medical Center, Parkland Health and Hospital Systems and Children's Medical Center. Informed consent was obtained for all infants in the study who had follow-up at the Thrive clinic.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Agarwal PK, Shi L, Rajadurai VS, Zheng Q, Yang PH, Khoo PC, et al. Factors affecting neurodevelopmental outcome at 2 years in very preterm infants below 1250 grams: a prospective study. J Perinatol. 2018;38:1093–100.
2. Belfort MB, Rifas-Shiman SL, Sullivan T, Collins CT, McPhee AJ, Ryan P, et al. Infant growth before and after term: effects on neurodevelopment in preterm infants. Pediatrics. 2011;128: e899–906.
3. Asztalos EV, Church PT, Riley P, Fajardo C, Shah PS, Canadian Neonatal N, et al. Neonatal factors associated with a good neurodevelopmental outcome in very preterm infants. Am J Perinatol. 2017;34:388–96.
4. Bauer SE, Schneider L, Lynch SK, Malleske DT, Shepherd EG, Nelin LD. Factors associated with neurodevelopmental impairment in bronchopulmonary dysplasia. J Pediatr. 2020;218:22–7 e2.
5. Belfort MB, Kuban KC, O'Shea TM, Allred EN, Ehrenkranz RA, Engelke SC, et al. Weight status in the first 2 years of life and neurodevelopmental impairment in extremely low gestational age newborns. J Pediatr. 2016;168:30–5 e32.
6. Cormack BE, Harding JE, Miller SP, Bloomfield FH. The influence of early nutrition on brain growth and neurodevelopment in extremely preterm babies: a narrative review. Nutrients. 2019;11:2029.
7. Ehrenkranz RA, Dusick AM, Vohr BR, Wright LL, Wrage LA, Poole WK. Growth in the neonatal intensive care unit influences neurodevelopmental and growth outcomes of extremely low birth weight infants. Pediatrics. 2006;117:1253–61.
8. Georgieff MK, Ramel SE, Cusick SE. Nutritional influences on brain development. Acta Paediatr. 2018;107:1310–21.
9. Nash A, Dunn M, Asztalos E, Corey M, Mulvihill-Jory B, O'Connor DL. Pattern of growth of very low birth weight preterm infants, assessed using the WHO Growth Standards, is associated with neurodevelopment. Appl Physiol Nutr Metab. 2011;36:562–9.
10. Duncan AF, Heyne RJ, Morgan JS, Ahmad N, Rosenfeld CR. Elevated systolic blood pressure in preterm very-low-birth-weight infants ≤ 3 years of life. Pediatr Nephrol. 2011;26:1115–21.
11. Frankfurt JA, Duncan AF, Heyne RJ, Rosenfeld CR. Renal function and systolic blood pressure in very-low-birth-weight infants 1-3 years of age. Pediatr Nephrol. 2012;27:2285–91.
12. Duncan AF, Frankfurt JA, Heyne RJ, Rosenfeld CR. Biomarkers of adiposity are elevated in preterm very-low-birth-weight infants at 1, 2 and 3 Y of age. Pediadtr Res. 2017;81:780–6.
13. Wickland J, Heyne R, Brown LS, Turer CB, Rosenfeld CR. Preterm very-low-birthweight (PT-VLBW) infants continue to exhibit high systolic blood pressure (SBP) and altered renal function at preadolescence. Presentation at AAP 94th Perinatal and developmental medicine symposium, Snow Mass, CO, June, 2019.
14. Meyers JM, Tan S, Bell EF, Duncan AF, Guillet R, Stoll BJ, et al. Eunice Kennedy Shriver National Institute of Child Health and Human Development Neonatal Research Network. Neurodevelopmental outcomes among extremely premature infants with linear growth restriction. J Perinatol. 2019; 39:193–202.
15. Belfort MB, Gillman MW, Buka SL, Casey PH, McCormick MC. Preterm infant linear growth and adiposity gain: trade-offs for later weight status and intelligence quotient. J Pediatr. 2013;163:1564–9 e2.
16. Rozé JC, Darmaun D, Boquien CY, Flamant C, Picaud JC, Savagner C, et al. The apparent breastfeeding paradox in very preterm infants: relationship between breast feeding, early weight gain and neurodevelopment based on results from two cohorts, EPIPAGE and LIFT. BMJ Open. 2012;2:e000834.
17. Franz AR, Pohlandt F, Bode H, Mihatsch WA, Sander S, Kron M, et al. Intrauterine, early neonatal, and postdischarge growth and neurodevelopmental outcome at 5.4 years in extremely preterm infants after intensive neonatal nutritional support. Pediatrics. 2009;123:e101–9.
18. Neubauer V, Fuchs T, Griesmaier E, Kager K, Pupp-Peglow U, Kiechl-Kohlendorfer U. Poor postdischarge head growth is related to a 10% lower intelligence quotient in very preterm infants at the chronological age of five years. Acta Paediatr. 2016;105:501–7.
19. Neubauer V, Griesmaier E, Pehböck-Walser N, Pupp-Peglow U, Kiechl-Kohlendorfer U. Poor postnatal head growth in very preterm infants is associated with impaired neurodevelopment outcome. Acta Paediatr. 2013;102:883–8.
20. Pavageau L, Whitham J, Rosenfeld CR, Heyne RJ, Brown LS, Brion LP. Valid serial length measurements in preterm infants permit characterization of growth patterns. J Perinatol. 2018;38:1694–70.
21. Brion LP, Rosenfeld CR, Heyne RJ, Brown SL, Lair CS, Burchfield PJ, et al. Adjustable feedings plus accurate serial length measurements decrease discharge weight-length disproportion in very preterm infants: quality improvement project. J Perinatol. 2019;39:1131–9.
22. Brion LP, Heyne R, Brown LS, Lair CS, Edwards A, Burchfield PJ, et al. Zinc deficiency limiting head growth to discharge in extremely low gestational age infants with insufficient linear growth: a cohort study. J Perinatol. 2020;40:1694–704.
23. Brion LP, Rosenfeld CR, Heyne R, Brown LS, Lair CS, Heyne E, et al. Association of age of initiation and type and of complementary foods with body mass index and weight-length at twelve months of age in preterm infants. J Perinatol. 2020;40:1394–404.
24. Albers CA, Grieve AJ. Bayley scales of infant and toddler development, 3rd edition. J Psychoeduc Assess. 2007;25:180–90.
25. Olsen IE, Groveman SA, Lawson ML, Clark RH, Zemel BS. New intrauterine growth curves based on United States data. Pediatrics 2010;125:e214–24.
26. Jobe AH, Bancalari E. Bronchopulmonary dysplasia. Am J Respir Crit Care Med. 2001;163:1723–9.
27. Papile LA, Burstein J, Burstein R, Koffler H. Incidence and evolution of subependymal and intraventricular hemorrhage: a study of infants with birth weights less than 1,500 gm. J Pediatr. 1978;92:529–34.
28. Kliegman RM, Walsh MC. Neonatal necrotizing enterocolitis: pathogenesis, classification, and spectrum of illness. Curr Probl Pediatr. 1987;17:213–88.
29. WHO Multicentre Growth Reference Study Group. WHO Child Growth Standards based on length/height, weight and age. Acta Paediatr. 2006;450:76.
30. Olsen IE, Lawson ML, Ferguson AN, Cantrell R, Grabich SC, Zemel BS, et al. BMI curves for preterm infants. Pediatrics. 2015;135:e572–81.
31. Albers CA, Grieve AJ. Test Review: Bayley, N. Bayley Scales of Infant and Toddler Development–Third Edition. San Antonio, TX: Harcourt Assessment. J Psychoeducational Assess. 2007;25:180–90.

32. Duncan AF, Bann C, Boatman C, Hintz SR, Vaucher YE, Vohr BR, et al. Eunice Kennedy Shriver National Institute of Child Health and Human Development Neonatal Research Network. Do currently recommended Bayley-III cutoffs overestimate motor impairment in infants born <27 weeks gestation? J Perinatol. 2015;35:516–21.

33. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol. 1996;49:1373–9.

34. Schreiber-Gregory D. Multicollinearity: what is it, why should we care, and how can it be controlled? Paper 1404–2017. http://support.sas.com/resources/papers/proceedings17/1404-2017.pdf. Accessed 20 Feb 2021.

35. Underwood MA. Human milk for the premature infant. Pediatr Clin North Am. 2013;60:189–207.

36. American Academy of Pediatrics Section on Breastfeeding. Breastfeeding and the use of human milk. Pediatrics 2012;129:e827–41.

37. Patra K, Hamilton M, Johnson TJ, Greene M, Dabrowski E, Meier PP, et al. NICU human milk dose and 20-month neurodevelopmental outcome in very low birth weight infants. Neonatology. 2017;112:330–6.

38. Fenton TR, Cormack B, Goldberg D, Nasser R, Alshaikh B, Eliasziw M, et al. "Extrauterine growth restriction" and "postnatal growth failure" are misnomers for preterm infants. J Perinatol. 2020;40:704–14.

39. Simon L, Théveniaut C, Flamant C, Frondas-Chauty A, Darmaun D, Rozé JC. In preterm infants, length growth below expected growth during hospital stay predicts poor neurodevelopment at 2 years. Neonatology. 2018;114:135–41.

40. Ramel SE, Demerath EW, Gray HL, Younge N, Boys C, Georgieff MK. The relationship of poor linear growth velocity with neonatal illness and two-year neurodevelopment in preterm infants. Neonatology. 2012;102:19–24.

41. Brion LP, Heyne R, Lair CS. Role of zinc in neonatal growth and brain growth: review and scoping review. Pediatr Res. 2020. https://doi.org/10.1038/s41390-020-01181-z. correction https://doi.org/10.1038/s41390-021-01425-6.