Factors Affecting Neonatal Weight Change One Month From Birth: A Retrospective Observational Study in rural Gadchiroli, India

Introduction:

Objectives:
Determine the factors affecting the change in the weight of neonates in the 28 days from birth.

Setting:
Data collected from SEARCH HBNC program, from April 1996 to March 2020

Variables:
Outcomes:
1. Absolute change in neonate weight from day 1 to day 28 ($wt_{28} - wt_1$)
2. Linear weight growth velocity, defined as change in weight per day divided by birth weight ($GV = \frac{wt_{28}-wt_1}{28} / wt_1$). This formula is used occasionally in the literature, and has the effect of converting the absolute change in weight, to a change in weight per kilogram of birth weight of a neonate. This is a more standardized metric, perhaps more relevant in a clinical setting
3. Exponential weight growth velocity, based on an exponential model in the literature from Patel et. al ($GV = 1000 * \frac{\log\left(\frac{wt_{28}}{wt_1}\right)}{28}$). This formula is used frequently in the literature, and is shown to give good empirical results.

There is substantial disagreement in the literature of weight gain of preterm/low birth weight babies on how to compute and report growth velocity.
- Patel and others measure neonate weights daily and show that the exponential model has excellent agreement empirically.
- Simon et al argues that weights standardized by z-scores, using existing empirical growth charts, are a better metric.
- Fenton argues for an alternative linear model, with growth velocity computed as $GV = \frac{wt_{28}-wt_1}{28} / \frac{wt_{28}+wt_1}{2}$,
- A systematic review of 151 studies shows considerable variation in the method used.

I have presented results for the 3 mentioned outcome variables. Inferences about which variables are significant are generally the same across choice of outcome, although there are some exceptions which have been pointed out.

Predictors:
Data for each infant contained a list of comorbidities, demographic information about the parents, and information about the birth.
- Demographic
  - Mother's education

    - Father's education
    - Maternal parity
    - Gestational age
    - Year of birth
    - Neonate sex
- Comorbidities
    - Previous Stillbirth
    - Hypothermia
    - Mother Feeding Problem
    - Baby Feeding Problem
    - Delayed breastfeeding
    - Twins
    - Birth asphyxia
    - Neonatal sepsis
    - Umbilical sepsis
    - Unexplained fever
    - Jaundice
    - Pneumonia
    - Congenital anomaly
    - Conjunctivitis
    - Bad obstetric history
    - Tobacco use in pregnancy
- Birth information
    - Location of delivery
    - Number of CHW visits,
    - CHW present at delivery

If a comorbidity was recorded as "missing" in the data, I have assumed that to mean the absence of the comorbidity. Other missing data is handled by imputation, as described later.

Comorbidities with extremely low prevalence were excluded from the analysis, which in this case were meconium aspiration and hyaline membrane disease.

Statistical methods:

Summary statistics for all variables were produced in Stata.

Crude analysis of association between predictors and outcome variables were done using independent two-sample two-sided t-tests for binary predictors, one-way Analysis of Variance (ANOVA) for categorical predictors with 3 or more groups, and ordinary linear regression for continuous predictors.

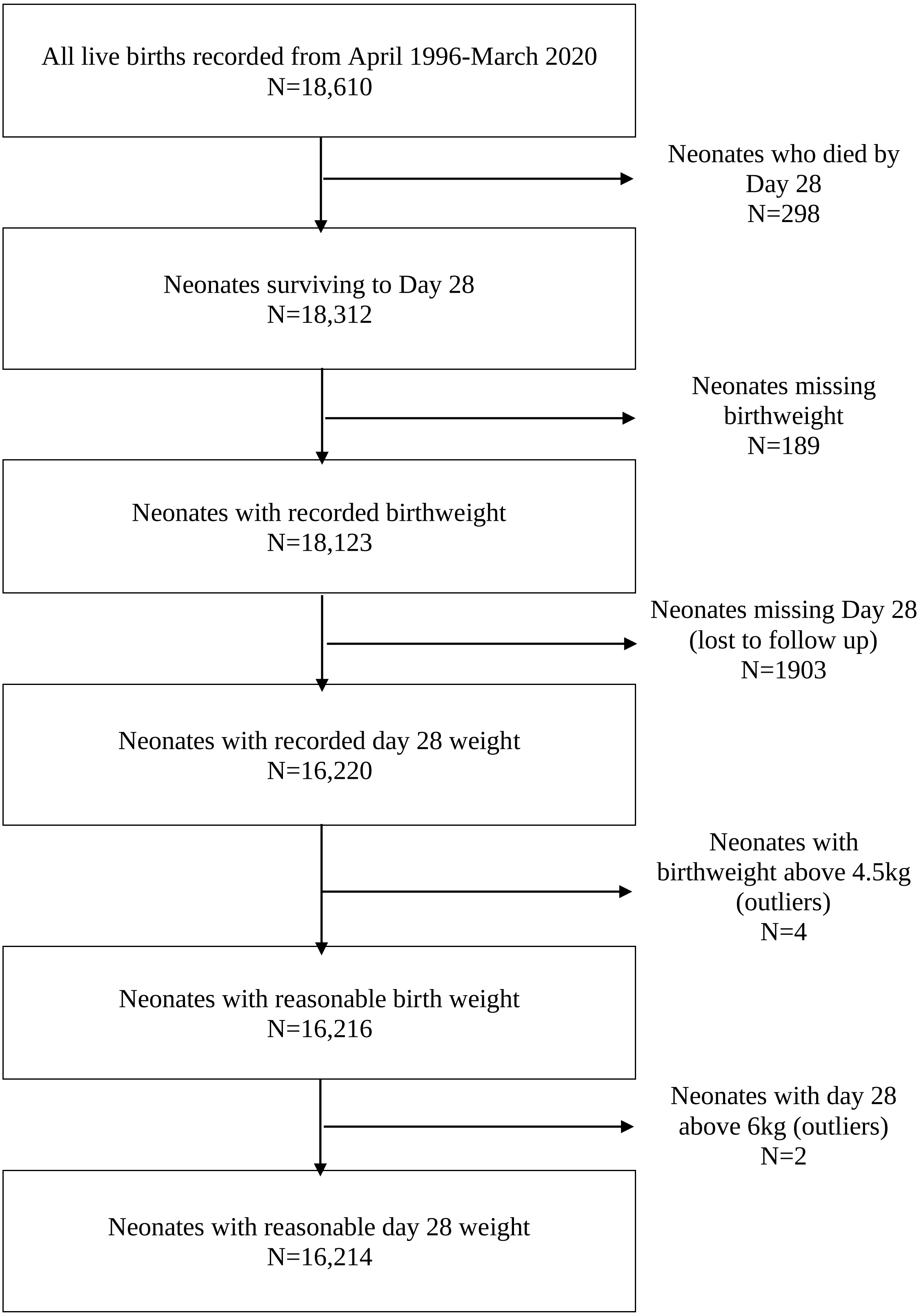
Multiple imputation was used to handle missing predictor data, with the assumption that data are Missing at Random (MAR). There was no missing outcome data, since all neonates without weight recorded at birth and follow-up were excluded. Multiple imputation was carried out using MICE (multiple imputation by chained equations) in Stata.

Multivariate analysis was conducted, using multiple linear regression to determine coefficients and significance of predictors. Parameter estimates with robust standard errors were used.

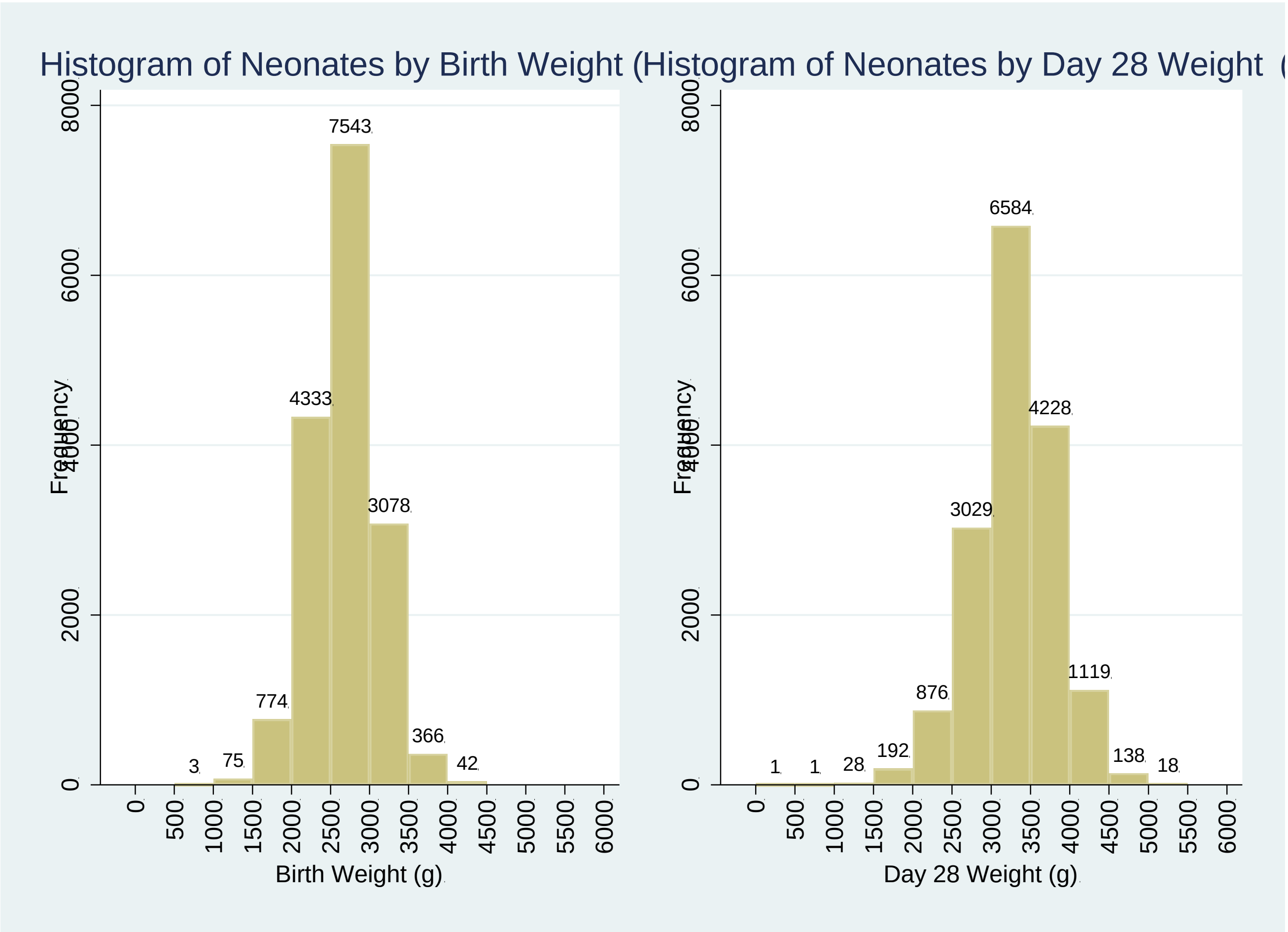
Multivariate regression was verified for usual conditions of multicollinearity, normality of residuals, and heteroscedasticity. Detected heteroscedasticity was handled by computing robust standard errors.

All analysis was performed in Stata.

Study Participant Flow Chart



The complete sample of live births recorded from April 1996 to March 2020 contained 18,610 neonates. 298 neonates died by day 28, and were excluded.
Of the surviving 18,312 neonates, 189 were missing a recording of birth weight. For home deliveries (n=9246), this could be because a CHW was not present to record the weight (n=37) or was present but did not record the weight (n=28). For institutional deliveries (n=9066), this could be because the health institution (hospital/PHC) did not record the birth weight, or because any recorded birth weight was not available to the CHW (n=124). Neonates with missing birth weights were excluded.
Of the remaining 18,123 surviving neonates with birthweights recorded, a further 1903 did not have weight recorded at day 28. This loss to follow-up could be because a mother moved to another village, or otherwise could not be found by a CHW.
With 16,220 neonates remaining, 6 outliers were excluded. Neonates with birthweight above 4 kg, an unreasonably high figure that could only be due to measurement or recording error, were excluded (n=4). Similarly, neonates with day 28 weight above 6 kg were excluded (n=2).
The final sample remaining and used for analysis thus contained 16,214 surviving, live-born neonates with reasonable, recorded weights at day 1 and day 28.



A further set of neonates could have been excluded as outliers, with a reasonable assumption of impossible weight recordings. The number of such neonates is extremely small, and so they are unlikely to affect the study outcomes in any meaningful way. Nevertheless, histograms of both birth and day 28 weights are presented below, on which further outlier exclusion can be defined.

Univariate Statistics

Summary statistics for all used variables are provided below.

Summary statistics for continuous variables are presented in Table 1. Gestational age ranged from 27.1 weeks to 55.6 weeks, with a mean of 39.5 weeks and a very low standard deviation of 2 weeks. The mean birth weight was 2.6 kg, and the mean day 28 weight was 3.2 kg, for a mean change in weight of 623 g. The linear weight growth velocity had a mean of 8.71 g/kg/day. In words, this means that on any given day, a neonate in this sample could be expected to gain an average of 8.71 g per kg of body weight on that day. Similarly, the exponential growth velocity had a mean of 7.62 g/kg/day.

Table 1. Summary statistics of continuous variables

Variable	N	Mean	SD	Min	Max
Gestational Age (weeks)	15746	39.45	1.96	27.09	55.61
Weight on day 1 (g)	16214	2615.88	416.45	750.00	4500.00
Weight on day 7 (g)	13950	2632.15	410.42	200.00	8752.00
Weight on day 15 (g)	15587	2849.42	455.87	1000.00	9000.00
Weight on day 21 (g)	5489	2833.39	541.22	800.00	9500.00
Weight on day 28 (g)	16214	3239.48	518.12	140.00	5275.00
Change in Weight (1-7) (g)	13950	30.48	152.66	-3050.00	5977.00
Change in Weight (7-15) (g)	13757	219.77	177.91	-5252.00	6750.00
Change in Weight (15-21) (g)	5277	173.48	167.25	-1000.00	5100.00
Change in Weight (21-28) (g)	5489	209.33	196.75	-6775.00	1300.00
Change in Weight (1-28) (g)	16214	623.61	296.40	-1250.00	2625.00
Linear growth velocity (g/kg/day)	16214	8.71	4.47	-31.87	125.00
Exponential growth velocity (g/kg/day)	16214	7.62	3.58	-79.59	53.72

Summary statistics for categorical variables are presented in Table 2. The sample is well-balanced by sex, at nearly 50% each male and female. The modal education level for mothers and fathers was 5th-10th class. The majority of mothers had gravidity of 1 or 2, and parity of 0 or 1. Roughly half of deliveries took place in both homes and institutions. 28% of neonates were twins. CHWs were present for 60.4% of deliveries. The modal number of CHW visits was 5-10, at 52%.

20% of neonates were identified to be at high risk on day 1, indicating that they were preterm, of low birth weight (<1 kg), and had a feeding problem on day 1 (high risk on day 1 is an interaction variable of those 3 conditions). Prevalence of feeding problems was relatively high, at 6.8% for babies and 3.8% for mothers. Just 0.5% of mothers delayed breastfeeding.

Comorbidities present included bad obstetric history (12%), Birth asphyxia (7.6%), neonatal sepsis (6%), unexplained fever (3.6%), hypothermia (2.8%), bacterial skin infections (2%), jaundice (1.3%), and umbilical sepsis (1%). The majority of neonates had no morbidities (69%), while 24% had just 1, and the remainder had multiple.

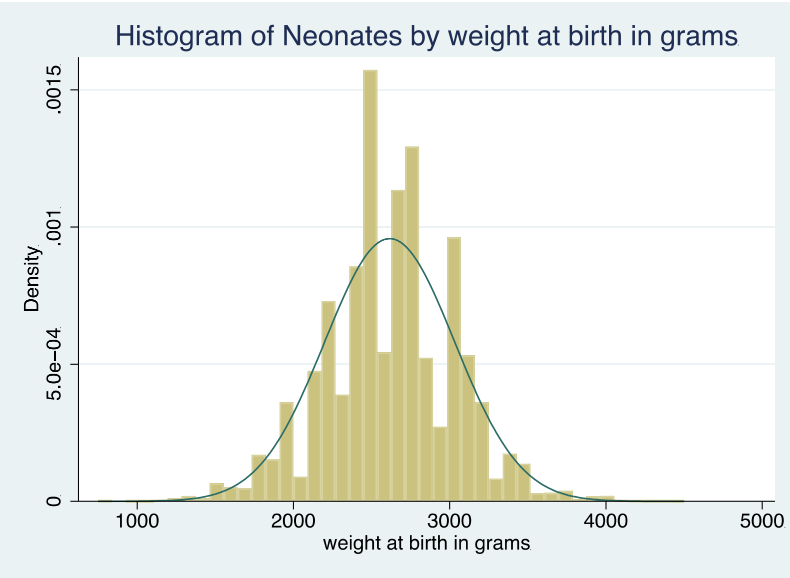
Table 2. Summary statistics of categorical variables

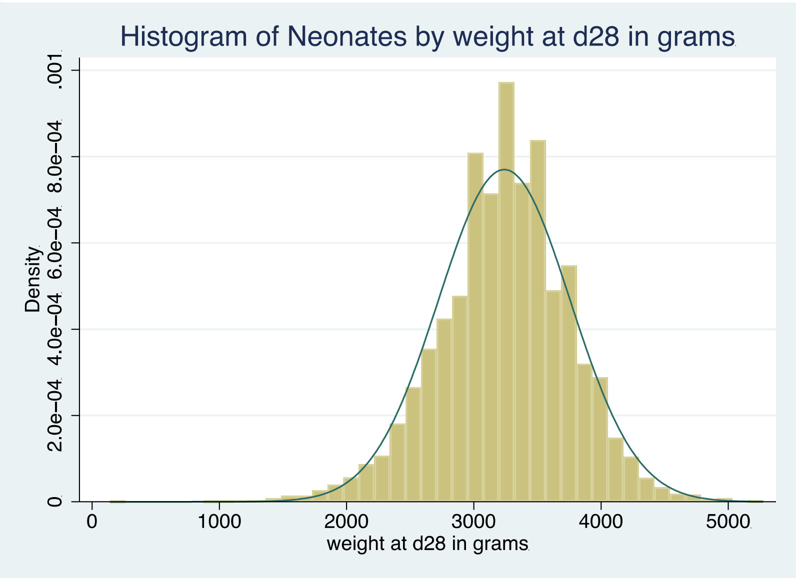
Variable	N	%
Sex		
Male	8195	50.54
Female	7978	49.20
Missing	41	0.25
Maternal education		
Illiterate	2374	14.64
1-4/Literate	771	4.76
5-10 class	8393	51.76
>=11 class	3961	24.43
Missing	715	4.41
Paternal education		
Illiterate	1755	10.82
1-4/Literate	1354	8.35
5-10 class	7488	46.18
>=11 class	4877	30.08
Missing	740	4.56
Gravidity		
1st	6431	39.66
2nd	5972	36.83
3rd	2184	13.47
>=4th	921	5.68
Missing	706	4.35
Parity		
0 child	7390	45.58
1 child	6086	37.54
2 child	1965	12.12
3 child	539	3.32
>=4 child	234	1.44
Institutional delivery		
Home	7939	48.96
Hospital/PHC	8274	51.03
Missing	1	0.01
Twins		
No	11751	72.47

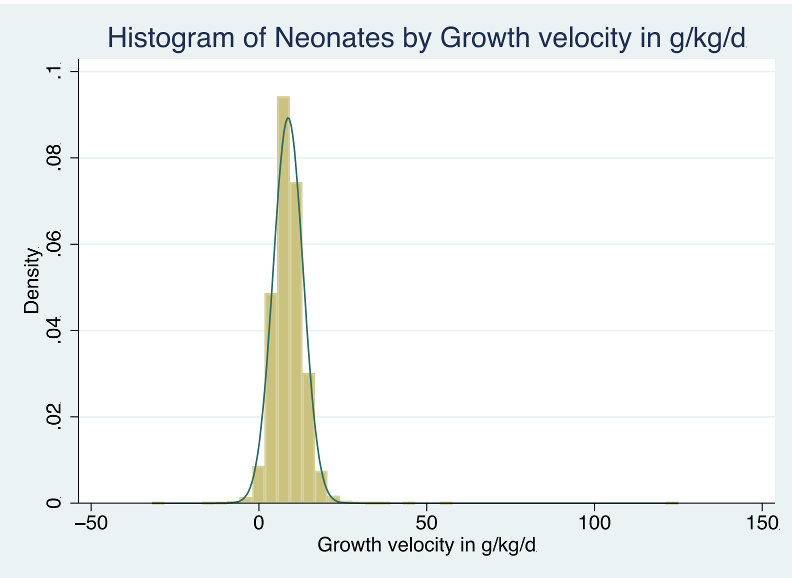
Yes	4463	27.53
High risk on day 1		
No	13022	80.31
Yes	3192	19.69
Feeding Problem (Baby)		
No	15112	93.20
Yes	1102	6.80
Feeding Problem (Mother)		
No	15600	96.21
Yes	614	3.79
Birth Asphyxia		
No	14978	92.38
Yes	1236	7.62
Congenital anomaly		
No	16152	99.62
Yes	62	0.38
Umbilical sepsis		
No	16029	98.86
Yes	185	1.14
Conjunctivitis		
No	16104	99.32
Yes	110	0.68
Unexplained fever		
No	15634	96.42
Yes	580	3.58
Hemorrhage		
No	16162	99.68
Yes	52	0.32
Hypothermia		
No	15766	97.24
Yes	448	2.76
Failure to gain weight		
No	14424	88.96
Yes	1790	11.04
Neonatal sepsis		
No	15242	94.01
Yes	972	5.99

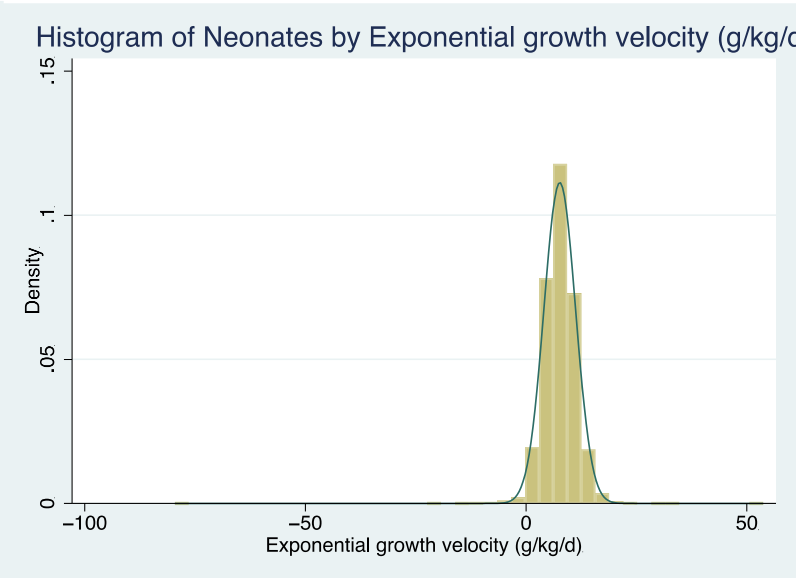
Pneumonia only		
No	15920	98.19
Yes	294	1.81
Bacterial skin infection		
No	15881	97.95
Yes	333	2.05
Jaundice		
No	16009	98.74
Yes	205	1.26
Bad obstetric history		
No	14283	88.09
Yes	1931	11.91
Tobacco use in pregnancy		
No	13528	83.43
Yes	2686	16.57
Program year		
1996-97	283	1.75
1997-98	444	2.74
1998-99	620	3.82
1999-00	840	5.18
2000-01	765	4.72
2001-02	697	4.30
2002-03	676	4.17
2003-04	640	3.95
2004-05	641	3.95
2005-06	694	4.28
2006-07	781	4.82
2007-08	617	3.81
2008-09	692	4.27
2009-10	715	4.41
2010-11	748	4.61
2011-12	719	4.43
2012-13	727	4.48
2013-14	725	4.47
2014-15	750	4.63
2015-16	769	4.74
2016-17	631	3.89

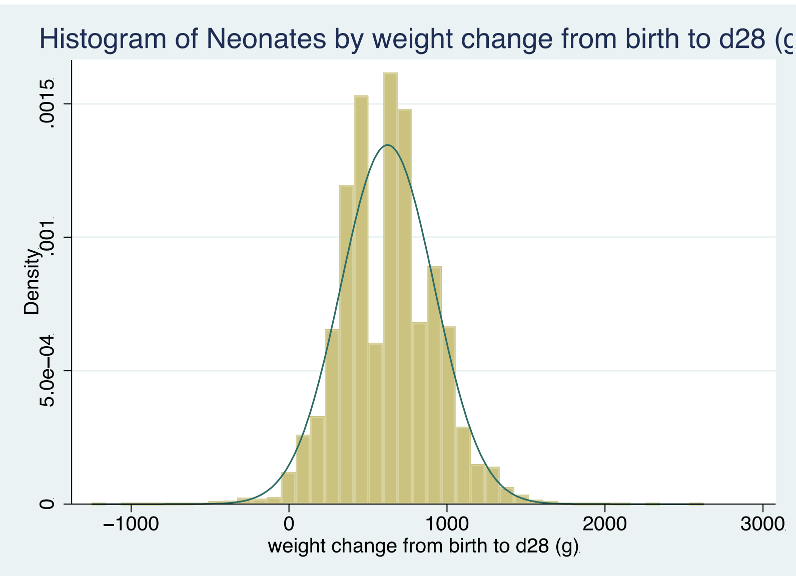
2017-18	731	4.51
2018-19	689	4.25
2019-20	620	3.82
Number of CHW visits		
0 visits	2770	17.08
1-5 visits	3745	23.10
5-10 visits	8457	52.16
10+ visits	1242	7.66
CHW Present at Delivery		
No	6419	39.59
Yes	9795	60.41
Delayed Breast Feeding		
No	16140	99.54
Yes	74	0.46
Number of comorbidities		
0	11207	69.12
1	3882	23.94
2	883	5.45
3	210	1.30
4	30	0.19
5	2	0.01

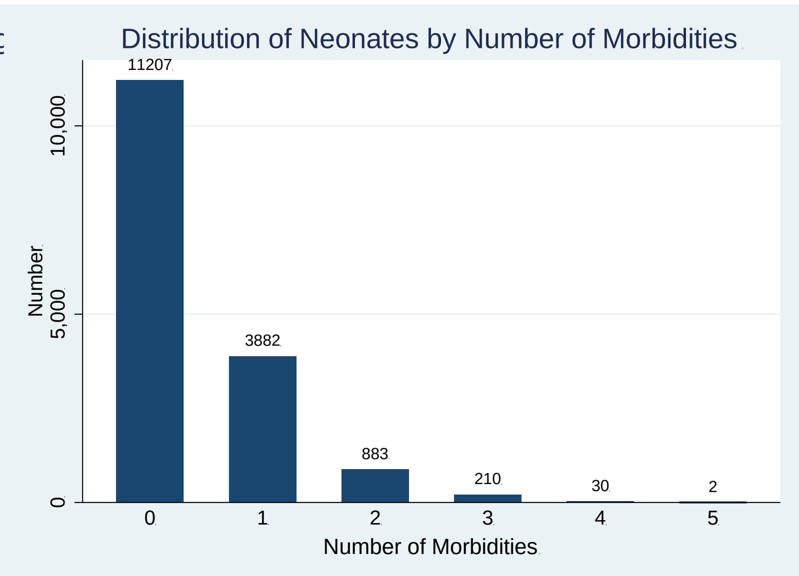

Histogram of Neonates by weight at birth in grams
Density
0
5.0e−04
.001
.0015
1000
2000
3000
4000
5000
weight at birth in grams


Histogram of Neonates by weight at d28 in grams
Density
0
2.0e−04
4.0e−04
6.0e−04
8.0e−04
.001
0
1000
2000
3000
4000
5000
weight at d28 in grams


Histogram of Neonates by Growth velocity in g/kg/d
Density
0
.02
.04
.06
.08
.1
−50
0
50
100
150
Growth velocity in g/kg/d


Histogram of Neonates by Exponential growth velocity (g/kg/d
Density
0
.05
.1
.15
−100
−50
0
50
Exponential growth velocity (g/kg/d)


Histogram of Neonates by weight change from birth to d28 (g
Density
0
5.0e−04
.001
.0015
−1000
0
1000
2000
3000
weight change from birth to d28 (g)


Distribution of Neonates by Number of Morbidities
Number
0
5,000
10,000
11207
3882
883
210
30
2
0
1
2
3
4
5
Number of Morbidities

Bivariate Statistics

Unadjusted bivariate associations between predictor variables and outcomes are presented below.

Unadjusted correlations between continuous predictor variables and all 3 outcomes are shown in Table 3. Gestational age was significantly positively correlated with absolute weight change. Neonates gained 15.6 g of weight from day 1 to 28 for each additional week of gestation.

Year of birth predicted absolute growth as well as growth velocity, with neonates gaining 5.5 g of weight from day 1 to 28 for each successive year after 1996 that they were born, or 60 milligrams per kg of body weight per day for the linear growth velocity, or 50 milligrams per kg per day for exponential growth velocity.

Birth weight was significantly negatively correlated with both linear and exponential metrics of growth velocity, likely due to improving neonatal care as well as improvements in economic development and healthcare over time. With the linear method, each additional gram of birth weight corresponded to a loss in 3 milligrams of weight per kg of body weight per day, indicating the existence of "catch-up" growth for low birthweight infants. The corresponding number for the exponential case was 2 milligrams per kg per day.

Table 3. Unadjusted Correlation between Outcomes and Continuous Predictors using Ordinary Linear Regression

Weight Change (g)			
Variable	**Coefficient**	**SE**	**p**
Gestational Age (weeks)	15.566***	1.197	0.000
Birth weight (g)	0.021	0.006	0.950
Year of Birth	5.473***	0.346	0.000
Linear Growth Velocity (g/kg/day)			
Variable	**Coefficient**	**SE**	**p**
Gestational Age (weeks)	-0.005	0.018	0.771
Birth weight (g)	-0.003***	809.000	0.000
Year of Birth	0.061***	0.005	0.000
Exponential Growth Velocity (g/kg/day)			
Variable	**Coefficient**	**SE**	**p**
Gestational Age (weeks)	0.005	0.014	0.727
Birth weight (g)	-0.002***	0.000	0.000
Year of Birth	.0502503***	0.004	0.000

Unadjusted correlations between binary predictor variables and all 3 outcomes are shown in Table 4. Several comorbidities were significantly associated with neonatal weight change and velocity. Neonates with birth asphyxia had a weight gain reduction of 17.2 g from day 1 to 28. Neonates born to mothers with bad obstetric history had a weight gain reduction of 35 g from day 1 to 28, and had reduced growth velocities of 0.5 grams per kg of body weight per day with the linear method, or 0.4 grams per kg per day with the exponential method.

Similarly, neonates with congenital anomalies had a weight gain reduction of 100 g, at 1.38 g/kg/day (linear) or 1.4 g/kg/day (exponential). Similar conclusions can be drawn for hemorrhages, hypothermia, jaundice, neonatal sepsis, and twin birth. Other variables were insignificantly correlated to all 3 outcomes.

Feeding was a major predictor of reduced weight gain. Delayed breastfeeding led to a weight gain reduction of 93.6 g over 28 days, or 0.94 mg/kg/day (exponential). Cases without baby feeding problems gained 99 g more over 28 days, or 0.45 g/kg/day (linear) and 506 mg/kg/day (exponential), and without a maternal feeding problems gained 76 g more over 28 days, or 0.77 g/kg/day (linear) and 0.72 g/kg/day (exponential).

Table 4. Unadjusted Correlation between Outcomes and Binary Predictors Using Student's t-tests

Variable	Mean Weight Change, Day 1-28 (g)	Mean (linear) Growth Velocity (g/kg/day)	Mean (exponential) Growth Velocity (g/kg/day)
Birth Asphyxia			
No	624.9184	8.702508	7.615298
Yes	607.7225	8.80136	7.636304
Difference	17.19592*	-.0988519	-.0210065
SE	(8.771042)	(.1321479)	(.1060538)
Bad Obstetric History			
No	627.8009	8.768575	7.664069
Yes	592.5909	8.277108	7.267998
Difference	35.21***	.4914671***	.396071***
SE	(7.181518)	(.1081999)	(.0868327)
Bacterial Skin Infection			
No	623.7783	8.710668	7.617981
Yes	615.4655	8.680273	7.565323
Difference	8.312823	.0303951	.0526575
SE	(16.41249)	(.2472536)	(.198427)
Congenital Anomaly			
No	623.9907	8.71535	7.622263
Yes	523.7903	7.327506	6.219446
Difference	100.2004**	1.387845*	1.402817**
SE	(37.70815)	(.5680872)	(.4558553)
Conjunctivitis			

No	623.2103	8.703909	7.612236
Yes	681.7727	9.608134	8.299646
Difference	-58.56247*	-.9042248*	-.6874108*
SE	(28.35427)	(.4271503)	(.3428039)
Breastfeeding Delayed			
No	624.0349	8.714595	7.621205
Yes	530.4054	7.717258	6.677764
Difference	93.62948**	.9973372	.9434407*
SE	(34.5281)	(.52022)	(.4174715)
Baby Feeding Problem			
No	630.3804	8.740778	7.65133
Yes	530.7305	8.288569	7.144748
Difference	99.64987***	.4522099**	.5065812***
SE	(9.215663)	(.1392871)	(.111747)
Mother Feeding Problem			
No	626.5066	8.739086	7.643978
Yes	549.9511	7.972159	6.928907
Difference	76.55546***	.7669272***	.7150712***
SE	(12.18047)	(.1836217)	(.1473334)
Sex			
Male	652.445	8.987458	7.830548
Female	593.2202	8.41562	7.390135
Difference	59.2248***	.5718377***	.4404134***
SE	(4.633638)	(.0700507)	(.056234)
Hemorrhage			
No	623.9449	8.714235	7.620389
Yes	518.75	7.407354	6.532249
Difference	105.1949*	1.306881*	1.08814*
SE	(41.16254)	(.6201478)	(.4976793)
High Risk on Day 1			
No	642.9125	8.705402	7.636379
Yes	544.8515	8.728981	7.537431
Difference	98.06103***	-.0235792	.0989481
SE	(5.803333)	(.0881927)	(.0707728)
Hypothermia			
No	626.7724	8.721562	7.633375
Yes	512.2321	8.304685	7.037092

Difference	114.5402***	.4168771	.5962829***
SE	(14.17314)	(.2139209)	(.1716333)
Institutional Delivery			
No	593.5533	8.346044	7.321741
Yes	652.4146	9.058847	7.899732
Difference	-58.86121***	-.7128034***	-.5779916***
SE	(4.633806)	(.0699305)	(.0561172)
CHW Present at Delivery			
No	618.174	8.670786	7.574375
Yes	627.1684	8.73577	7.644767
Difference	-8.994337	-.0649846	-.0703921
SE	(4.759432)	(.0717063)	(.0575449)
Jaundice			
No	622.9591	8.700979	7.609991
Yes	674.2488	9.417895	8.156368
Difference	-51.28969*	-.716916*	-.5463772*
SE	(20.83045)	(.3138159)	(.2518493)
Neonatal Sepsis			
No	627.7574	8.732164	7.641809
Yes	558.534	8.363173	7.226281
Difference	69.22343***	.3689909*	.4155289***
SE	(9.790755)	(.1476952)	(.118507)
Pneumonia Only			
No	623.0162	8.701322	7.60981
Yes	655.6293	9.182286	8.000777
Difference	-32.61305	-.4809637	-.3909668
SE	(17.44403)	(.262793)	(.2108976)
Tobacco During Pregnancy			
No	621.6786	8.675464	7.5867
Yes	633.3228	8.884201	7.768998
Difference	-11.64419	-.2087363*	-.1822984*
SE	(6.260697)	(.0943123)	(.075686)
Twin			
No	644.2273	8.927955	7.796698
Yes	569.3162	8.136286	7.143492
Difference	74.91114***	.7916694***	.6532063***
SE	(5.178489)	(.0782685)	(.0628012)

Umbilical Sepsis			
No	624.0822	8.709997	7.620011
Yes	582.4865	8.714068	7.347292
Difference	41.59568	-.0040712	.2727186
SE	(21.91551)	(.3301907)	(.2649779)
Unexplained Fever			
No	623.2183	8.704039	7.613762
Yes	634.1	8.871897	7.701456
Difference	-10.88169	-.1678576	-.0876941
SE	(12.53371)	(.1888184)	(.1515337)

Unadjusted correlations between ordinal predictor variables and all 3 outcomes are shown in Table 5.

[Note: These correlation tests were done using the ANOVA method, which is related to the Student's t-test. It is used to compare means among 3 or more groups. The null hypothesis, that the means of all groups are the same, is tested using an F-test, and a low p-value allows us to reject this hypothesis and conclude that the outcome variable is, in fact, significant different across the 3 groups.]

All outcomes were significantly different across number of CHW visits, although the results were somewhat unexpected. For absolute weight change, 10+ visits seemed to be worse than even 0 visits. Perhaps neonates/mothers who were already extremely unhealthy, or needed further medical care, were likely to have more CHW visits.

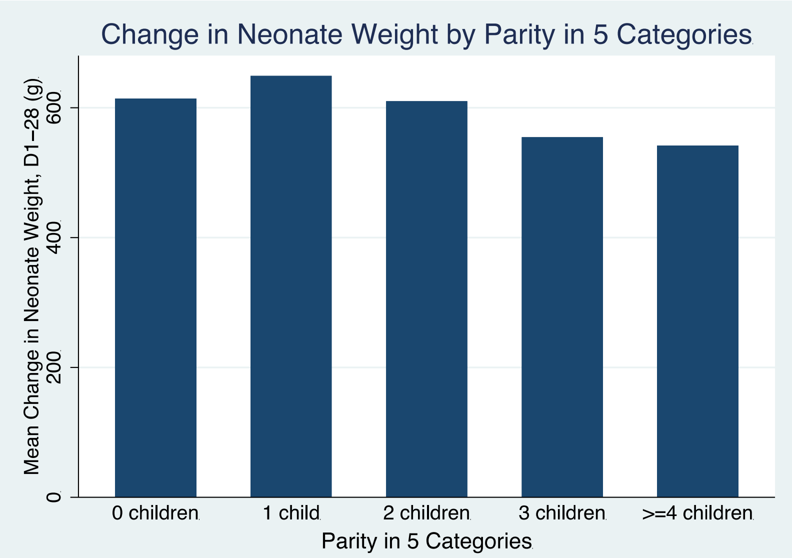
All 3 outcome variables showed a statistically significant difference across levels of both mother's and father's education, with higher education levels predicting higher weight gain, as well as growth velocity.

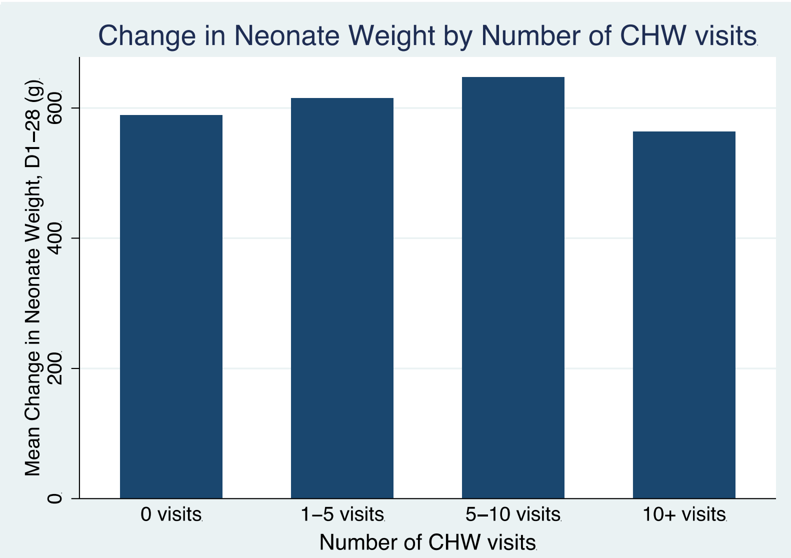
All outcomes were also significantly different across parity and gravidity, showing that mothers on their 2nd pregnancy (with 1 previous child) had the best-performing neonates.

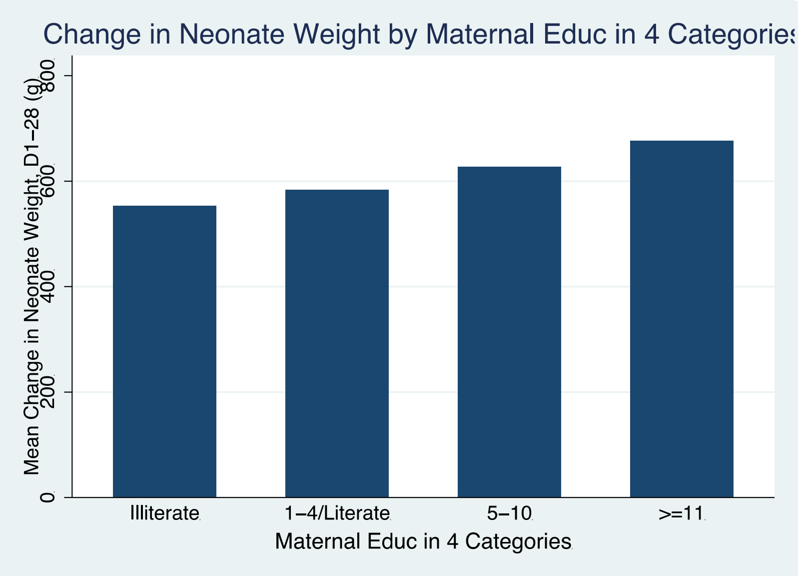
Table 5. Unadjusted Correlation between Outcomes and Categorical Predictors Using Analysis of Variance (ANOVA)

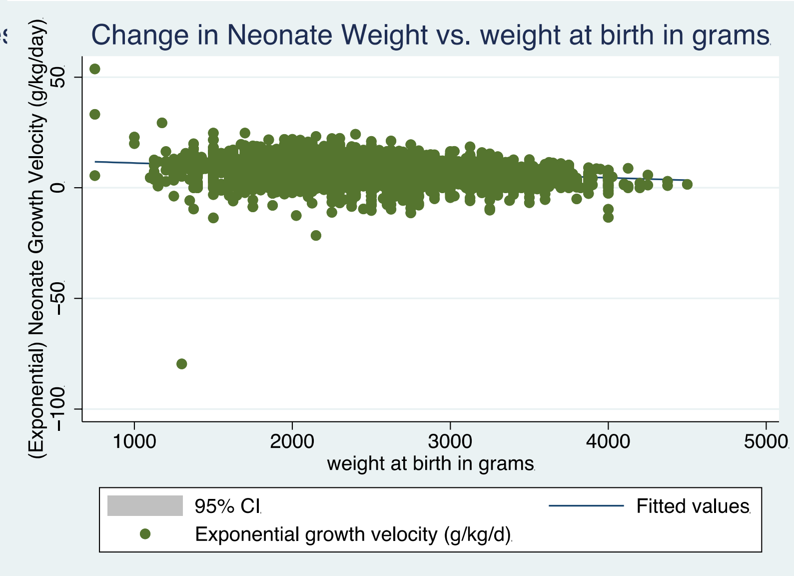
Variable	Weight Change	Linear Growth Velocity	Exponential Growth Velocity
No. of CHW visits			
0	589.0220	8.3760	7.3190
1	615.2020	8.2810	7.2780
2	647.4020	8.8620	7.7660
3	589.0224	9.7162	8.2871
p-value(ANOVA F test)	0.0000	0.0000	0.0000
Mother's Education			
Illiterate	553.4020	7.9750	7.0130
1-4/Literate	583.5110	8.2670	7.2630
5-10 class	626.7120	8.7700	7.6660
>=11	676.4500	9.1890	8.0080
p-value(ANOVA F test)	0.0000	0.0000	0.0000
Father's education			
Illiterate	561.8270	8.0220	7.0590
1-4/Literate	579.4980	8.2630	7.2590
5-10 class	628.7740	8.8380	7.7230
>=11	658.1550	8.9540	7.8090
p-value(ANOVA F test)	0.0000	0.0000	0.0000
Gravidity			
1st preg	618.1830	8.8600	7.7280
2nd preg	645.7740	8.8070	7.7010
3rd preg	623.0810	8.5700	7.5140
>=4 preg	559.1340	7.6970	6.8080
p-value(ANOVA F test)	0.0000	0.0000	0.0000
Parity			
0 children	613.7920	8.8000	7.6790
1 child	649.1220	8.8560	7.7420
2 children	610.1620	8.3520	7.3440
3 children	554.7350	7.6160	6.7520
>=4 children	541.5600	7.5990	6.6820
p-value(ANOVA F test)	0.0000	0.0000	0.0000

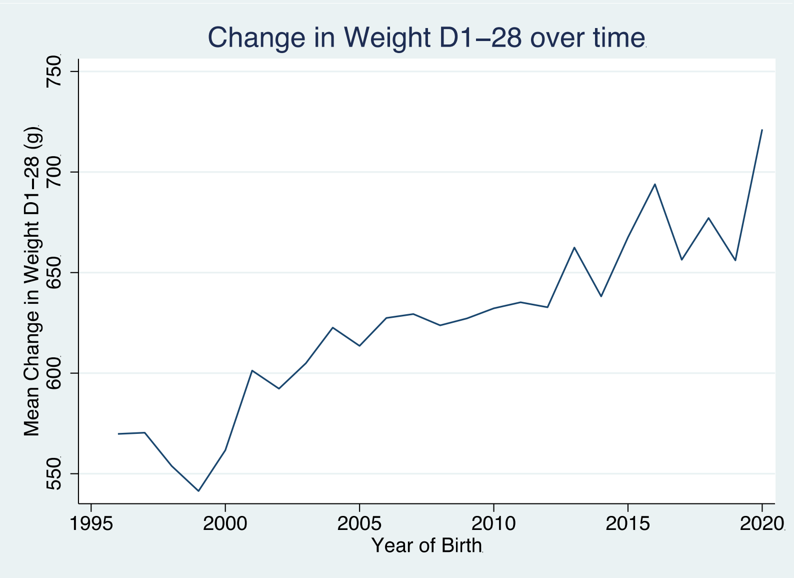
Selected bivariate charts are presented below. The difference in neonatal weight gain by parity, by number of CHW visits, and by maternal education, are presented for illustration. The linear correlation between growth velocity (exponential) and birth weight is shown, and it is a clearly negative linear trend. Absolute neonatal weight change over 28 days is shown from 1996-2020, with a clear positive trend.


Change in Neonate Weight by Parity in 5 Categories
Mean Change in Neonate Weight, D1–28 (g)
0
200
400
600
0 children
1 child
2 children
3 children
>=4 children
Parity in 5 Categories


Change in Neonate Weight by Number of CHW visits
Mean Change in Neonate Weight, D1–28 (g)
0
200
400
600
0 visits
1–5 visits
5–10 visits
10+ visits
Number of CHW visits


Change in Neonate Weight by Maternal Educ in 4 Categories
Mean Change in Neonate Weight, D1–28 (g)
0
200
400
600
800
Illiterate
1–4/Literate
5–10
>=11
Maternal Educ in 4 Categories


Change in Neonate Weight vs. weight at birth in grams
(Exponential) Neonate Growth Velocity (g/kg/day)
−100
−50
0
50
1000
2000
3000
4000
5000
weight at birth in grams
95% CI
Fitted values
Exponential growth velocity (g/kg/d)


Change in Weight D1–28 over time
Mean Change in Weight D1–28 (g)
550
600
650
700
750
1995
2000
2005
2010
2015
2020
Year of Birth

Multiple Imputation for Missing Data

Missing data present in the sample for predictor variables, is handled through a statistical procedure called multiple imputation.

Note: The method used here, Multiple Imputation by Chained Equations (MICE), uses a stepwise procedure to generate missing data, based on patterns in the existing data. It has been shown to be less biased than the alternative, of simply dropping all missing data (listwise deletion, also called complete case analysis).

In brief, MICE works to get estimates for each variable, by performing regression to obtain the predicted value of that variable based on the rest of the data set.

For example, with missing gestational age, a linear regression is performed with gestational age as the outcome and all other variables as the predictors. Using this regression, predicted values of gestational age are generated (along with an added randomization component). [For this regression, all other missing variables are temporarily replaced with the mean of all available data.]

This step is performed for all missing variables in turn, with a complete cycle called one imputation. The process is then repeated several times, and the resulting predictions of missing values are pooled together.

In this case, I have used 25 imputations. The type of regression performed for each missing variable is shown, with linear regression for continuous gestational age, logistic regression for the binary variables of institutional delivery and sex, and predictive mean matching for the ordinal variables of gravidity and parent education. [Predictive mean matching (PMM) is used in place of ordinal logistic regression to ensure convergence of the algorithm. PMM uses a k-nearest-neighbours approach to generate predictions, and here I have specified to use the 10 nearest-neighbours].

The command used in Stata is

```
mi impute chained (regress) gage (logit) neo pnm_ch institutional_delivery female
(pmm, knn(10)) gravida meduct heduct = twins ba bwt hrisk1 isfad delay feed_b
feed_m cong umbi conj jaun unfev hemo hypo bsi wt28_g wt1_g boh para tobause
delivery_date CHW_visits,add(25) rseed(12345) augment
```

Table 6. Handling of Missing Data by Multiple Imputation

Variable	Method	Complete	Incomplete	Imputed	Total
Gestational Age	Linear Regression	15746	468	468	16214
Institutional Delivery	Logistic Regression	16213	1	1	16214
Sex	Logistic Regression	16173	41	41	16214
Gravidity	Predictive Mean Matching	15508	706	706	16214
Mother's Education	Predictive Mean Matching	15499	715	715	16214

Father's Education	Predictive Mean Matching	15474	740	740	16214

Multiple Regression

Multiple regression is performed on all 3 outcome variables, including all relevant predictor variables.

With multiply imputed data, regression estimates are pooled from the 25 imputations performed by Stata's *mi estimate* command. Results are presented both with multiple imputation, and with the alternative of complete case analysis (using only observations without missing data).

Statistical tests in Stata indicate that the data may contain heteroskedasticity, which means that the residuals of the regression do not have the same variance across the dataset observations. This is a violation of an assumption of linear regression. To correct for this, we use Huber-White robust standard errors in the regression, to ensure that the standard errors and p-values are consistent.

A literature review showed that many authors consider gestational age to be a collider variable in the estimation of neonatal outcomes. Thus, regression tables are presented both with and without gestational age as a predictor.

A simpler alternative model is also presented with a *morbidity_num* variable, which combines all comorbidities into a single variable.

A regression table for absolute weight change as an outcome is presented in Table 7(a), for linear weight change velocity in Table 7(b), and for exponential weight change velocity in Table 7(c). I have provided interpretations in words for Tables 7(a) and 7(b), and the interpretation of Table 7(c) will be identical to Table 7(b) except for slightly different numbers. In addition, I will focus on model 4 in each case, which is the regression including gestational age as a predictor, and using multiple imputation as described above. Depending on which model(s) are chosen to present, this interpretation will be slightly modified.

Interpreting Table 7(a):

Birth weight is a highly significant ($p<0.001$) predictor of weight change, with a negative coefficient indicating the presence of "catch-up" growth for lower birthweight neonates, and corresponding "catch-down" growth for higher birthweight neonates. For a 100 g reduction in birthweight, a neonate will gain 5.8 g more in weight over 28 days.
Gestational age was found to be a significant predictor of weight gain ($p<0.001$), with each additional week of gestation leading to 15.3 g of additional weight gain over 28 days.
Year of birth is a significant ($p<0.01$) predictor of weight gain, with neonates born in later years gaining more weight. Each successive year gives neonates an additional 2.5 g of weight gain.
Sex is a highly significant ($p<0.001$) predictor of weight gain, with female neonates gaining substantially less weight (64.3 g less over 28 days) than males, despite their lower birthweight.
Neonates delivered at institutions have a significantly ($p<0.001$) higher weight gain of 51.6 g.
Neonates identified as high risk on day 1, as preterm, low birthweight, and having a feeding problem, gained significantly less weight (52.4 g less, $p<0.001$).
Neonates having a feeding problem on any of the 28 days gained significantly less weight (77.5g, $p<0.001$). Similarly, neonates whose mothers had feeding problems on any day

gained significantly less weight (30.2g, $p<0.05$) although the effect size and significance was less pronounced than for a neonate feeding problem.
The presence of a CHW at the delivery, perhaps indicating better access to and trust in health services, was strongly correlated with increased weight gain. Neonates delivered with CHWs present gained 27.7 g more over 28 days ($p<0.001$).
Delayed breastfeeding was a predictor of decreased weight gain, with neonates fed after a delay gaining 82.8 g less ($p<0.05$) than those fed immediately.
Of the range of comorbidities included, only hypothermia had a significant ($p<0.05$) effect on weight gain, of a 36.9 g reduction in weight gain over 28 days.
Bad obstetric history was significantly associated with lower weight gain (28 g, $p<0.001$). Strangely, tobacco usage during pregnancy was correlated with higher weight gain (20.4 g, $p<0.001$). This relationship is found in the literature as well, suggesting that this is a real effect.
The number of visits from a CHW was negatively correlated with weight gain, with cases involving 1-5 or 5-10 CHW visits having significantly less weight gain (65.7 g, $p<0.001$ and 19.1 g, $p<0.05$ respectively) than those with 0 CHW visits. More CHW visits may indicate a worse pre-existing health situation, such that neonates gaining weight more slowly requiring more CHW visits.
Education levels of both parents were correlated with weight gain. An increase of education level from Illiterate to Literate/Class 1-4 had no significant effect for either parent. A significant effect was found for an increase to either Class 5-10 or Class 11 and above for both parents. The effect size for the mother's education was significantly higher than that of the father's education, at both levels. Neonates born to mothers with education from Class 5-10 gained 25.9 g ($p<0.01$) more than those with illiterate mothers, and similarly to mothers with education from Class 11 and above gained 63.3 g ($p<0.001$) more. Similar effects were seen for father's education from Class 5-10 (20.9 g, $p<0.05$) and from Class 11 and above (37g, $p<0.001$).
Maternal parity was a significant predictor of increased weight gain. As compared to neonates born to nulliparous mothers, neonates born to mothers with 1 previous birth gained 44.7 g ($p<0.001$) more over 28 days, and similarly for those with 2 previous births. (38.9 g, $p<0.001$). Neonates born to mothers with 3 or more previous births had no weight gain advantage over those with nulliparous mothers.

Interpreting Table 7(b):

All coefficients in Table 7(b) are of growth velocities, measured in g/kg/day. This can be interpreted as the average daily change in weight of a neonate, per kilogram of its weight *on that day*.

Birth weight is a highly significant ($p<0.001$) predictor of weight growth velocity, with a negative coefficient indicating the presence of higher growth velocity for lower birthweight neonates, and corresponding lower growth for higher birthweight neonates. For a 100 g reduction in birthweight, a neonate's weight gain velocity will be reduced by 0.4 g/kg/day.
Gestational age was found to be a significant predictor of weight growth velocity ($p<0.001$), with each additional week of gestation leading to increased growth velocity of 0.24 g/kg/day.
Year of birth is *not* a significant ($p>0.01$) predictor of weight growth velocity, in contrast to the significant association with absolute weight change.
Sex is a highly significant ($p<0.001$) predictor of weight growth velocity, with female neonates having a lower weight velocity of 0.89 g/kg/day than males, despite their lower birthweight.

Neonates delivered at institutions gain weight significantly faster than those delivered at home (0.72 g/kg/day, p<0.001).
have a significantly (p<0.001) higher weight growth velocity of 51.6 g.
Neonates identified as high risk on day 1, as preterm, low birthweight, and having a feeding problem, had a significantly lower growth velocity (0.58 g/kg/day less, p<0.001).
Neonates having a feeding problem on any of the 28 days had a significantly lower growth velocity (1.23 g/kg/day, p<0.001). Similarly, neonates whose mothers had feeding problems on any day had a significantly lower growth velocity (0.62g/kg/day, p<0.01) although the effect size and significance was less pronounced than for a neonate feeding problem.
The presence of a CHW at the delivery, perhaps indicating better access to and trust in health services, was strongly correlated with higher growth velocity. Neonates delivered with CHWs present gained 0.25 g/kg/day more over 28 days (p<0.001).
Delayed breastfeeding was a predictor of decreased weight gain velocity, with neonates fed after a delay gaining 1.4 g/kg/day less (p<0.05) than those fed immediately.
None of the comorbidities included had any significant effect on growth velocity.
Bad obstetric history was significantly associated with lower weight gain velocity (0.41 g/kg/day, p<0.001).
Strangely, tobacco usage during pregnancy was correlated with higher weight gain velocity (0.28 g/kg/day, p<0.001). This relationship is found in the literature as well, suggesting that this is a real effect.
The number of visits from a CHW was negatively correlated with weight gain velocity, with cases involving 1-5 CHW visits having significantly lower weight gain velocity (0.9 g/kg/day, p<0.001). More CHW visits may indicate a worse pre-existing health situation, such that neonates gaining weight more slowly requiring more CHW visits.
Education levels of both parents were correlated with weight gain velocity. An increase of education level from Illiterate to Literate/Class 1-4 had no significant effect for either parent. A significant effect was found for an increase to either Class 5-10 or Class 11 and above for both parents. The effect size for the mother’s education was significantly higher than that of the father’s education, at both levels. Neonates born to mothers with education from Class 5-10 gained 0.44 g/kg/day (p<0.001) more than those with illiterate mothers, and similarly to mothers with education from Class 11 and above gained 0.94 g/kg/day (p<0.001) more. Similar effects were seen for father’s education from Class 5-10 (0.35 g/kg/day, p<0.01) and from Class 11 and above (0.55 g/kg/day, p<0.001).
Maternal parity was a significant predictor of increased weight gain velocity. As compared to neonates born to nulliparous mothers, neonates born to mothers with 1 previous birth gained 0.61 g/kg/day (p<0.001) more over 28 days, and similarly for those with 2 previous births. (0.51 g/kg/day, p<0.001). Neonates born to mothers with 3 or more previous births had no weight gain velocity advantage over those with nulliparous mothers.

Table 7(a)

Table 7(a). Multiple Linear Regression of Change in Birth Weight, Day 1-28 (g) on Predictors

Predictor	**Complete Case Analysis** ***Coefficient (SE)***	**Chained MI** ***Coefficient (SE)***	**Complete Case Analysis (controlling for Gestational Age)** ***Coefficient (SE)***	**Chained MI (controlling for Gestational Age)** ***Coefficient (SE)***	**Complete Case Analysis (combining comorbidities)** ***Coefficient (SE)***	**Chained MI (combining comorbidities)** ***Coefficient (SE)***
Birth Weight	-0.0427*** (0.00721)	-0.0396*** (0.00693)	-0.0577*** (0.00737)	-0.0528*** (0.00702)	-0.0402*** (0.00719)	-0.0374*** (0.00691)
Year of Birth	2.053** (0.734)	1.353 (0.704)	2.452** (0.761)	1.639* (0.703)	1.977** (0.730)	1.284 (0.700)
Female	-62.77*** (4.665)	-60.59*** (4.556)	-64.30*** (4.720)	-63.68*** (4.551)	-63.40*** (4.661)	-61.30*** (4.553)
Institutional Delivery	51.62*** (7.181)	53.76*** (6.979)	49.72*** (7.235)	51.61*** (6.956)	51.54*** (7.161)	53.74*** (6.960)
Twin	-0.429 (8.814)	-3.281 (8.446)	-2.752 (8.962)	-6.652 (8.419)	0.450 (8.791)	-2.572 (8.422)
High Risk on Day 1	-66.81*** (8.938)	-76.01*** (8.629)	-46.08*** (9.215)	-52.37*** (8.894)	-69.53*** (8.952)	-78.52*** (8.631)
Baby Feeding Problem	-69.91*** (14.53)	-74.83*** (14.02)	-74.08*** (14.66)	-77.54*** (13.98)	-65.42*** (11.37)	-69.96*** (11.05)
Mother Feeding Problem	-31.79* (13.89)	-26.42 (13.70)	-39.91** (13.82)	-32.65* (13.67)	-30.20* (13.98)	-25.34 (13.79)
CHW Present At Delivery	30.30*** (6.813)	29.92*** (6.647)	24.15*** (7.019)	27.73*** (6.627)	31.64*** (6.780)	31.19*** (6.613)
Delayed Breastfeeding	-70.96	-86.90*	-90.90*	-82.75*	-76.99	-91.87*

	(45.26)	(39.64)	(45.02)	(39.53)	(45.19)	(39.55)
Birth Asphyxia	-3.191 (9.054)	-1.586 (8.943)	-4.786 (9.038)	-3.993 (8.891)		
Congenital Anomaly	-79.41 (48.97)	-72.72 (48.52)	-84.47 (50.16)	-70.03 (49.43)		
Umbilical Sepsis	10.81 (24.65)	4.667 (24.37)	13.44 (24.78)	5.171 (24.38)		
Conjunctivitis	55.57 (32.59)	58.14 (32.47)	52.49 (31.68)	61.44 (32.60)		
Unexplained Fever	29.52* (14.32)	27.67* (14.00)	27.36 (14.36)	25.85 (13.95)		
Hemorrhage	-60.20 (38.91)	-58.86 (38.86)	-63.85 (39.37)	-63.69 (39.36)		
Hypothermia	-38.63* (17.02)	-37.92* (16.70)	-39.47* (17.03)	-36.89* (16.52)		
Neonatal Sepsis	2.851 (14.93)	4.518 (14.46)	8.121 (15.01)	8.523 (14.45)		
Pneumonia Only	22.82 (18.26)	19.72 (17.19)	23.51 (18.15)	20.91 (17.14)		
Bacterial Skin Infection	28.61 (17.16)	27.10 (16.78)	24.77 (16.76)	26.14 (16.70)		
Jaundice	30.19 (24.75)	29.28 (24.56)	21.19 (24.44)	28.25 (24.35)		
Bad Obstetric History	-32.24*** (7.836)	-30.33*** (7.821)	-31.87*** (7.934)	-28.01*** (7.809)		
Tobacco Use During Pregnancy	18.91** (6.265)	20.28** (6.236)	20.06** (6.283)	20.44*** (6.213)	18.06** (6.273)	19.46** (6.244)
Number of CHW visits						
1-5 visits	-59.48*** (10.75)	-68.67*** (10.51)	-60.64*** (10.96)	-65.71*** (10.46)	-60.97*** (10.77)	-69.90*** (10.53)
5-10 visits	-24.63** (8.504)	-27.80*** (8.336)	-16.48 (8.541)	-19.09* (8.352)	-24.19** (8.518)	-27.23** (8.348)
10+ visits	-47.44***	-38.16***	-22.78	-14.78	-46.08***	-36.86***

	(11.46)	(11.18)	(11.75)	(11.40)	(11.47)	(11.19)
Mother's Education						
1-4/Literate	15.05	13.75	18.24	16.19	15.54	14.29
	(12.12)	(12.00)	(12.18)	(11.96)	(12.18)	(12.05)
5-10 class	25.33**	25.45**	25.62**	25.88**	25.65**	25.80**
	(8.365)	(8.320)	(8.389)	(8.302)	(8.369)	(8.321)
>= 11 class	62.29***	62.25***	65.02***	63.30***	62.95***	62.90***
	(10.02)	(9.979)	(10.07)	(9.957)	(10.03)	(9.985)
Father's Education						
1-4/Literate	0.388	-0.711	0.903	0.154	2.953	1.793
	(10.47)	(10.46)	(10.55)	(10.43)	(10.45)	(10.44)
5-10 class	18.81*	18.95*	19.80*	20.86*	20.08*	20.16*
	(8.651)	(8.619)	(8.702)	(8.600)	(8.651)	(8.616)
>= 11 class	33.30***	33.71***	37.79***	36.96***	34.76***	35.10***
	(9.649)	(9.587)	(9.722)	(9.564)	(9.649)	(9.584)
Parity						
1 births	39.94***	44.80***	41.09***	44.73***	35.61***	40.55***
	(5.413)	(5.305)	(5.441)	(5.289)	(5.305)	(5.192)
2 births	34.78***	38.13***	38.31***	37.89***	24.70**	28.44***
	(8.222)	(8.157)	(8.295)	(8.156)	(7.895)	(7.819)
3 births	9.917	12.08	17.06	14.37	-5.555	-2.803
	(13.53)	(13.47)	(13.60)	(13.42)	(13.17)	(13.10)
>=4 births	24.01	21.66	27.48	21.48	5.068	3.627
	(22.87)	(22.74)	(23.49)	(22.74)	(22.41)	(22.25)
Gestational Age (weeks)			14.37***	15.26***		
			(1.421)	(1.391)		
Number of Comorbidities					-2.975	-2.310
					(4.129)	(4.075)
Intercept	-3420.4*	-2024.4	-4755.4**	-3175.6*	-3274.9*	-1890.1
	(1470.8)	(1411.5)	(1527.5)	(1412.0)	(1463.1)	(1403.9)
adj. R2	0.061	.	0.069	.	0.059	.
N	15422	16214	14970	16214	15422	16214

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 7(b)

Table 7(b). Multiple Linear Regression of Linear Growth Velocity (g/kg/day) on Predictors						
Predictor	**Complete Case Analysis** *Coefficient (SE)*	**Chained MI** *Coefficient (SE)*	**Complete Case Analysis (controlling for Gestational Age)** *Coefficient (SE)*	**Chained MI (controlling for Gestational Age)** *Coefficient (SE)*	**Complete Case Analysis (combining comorbidities)** *Coefficient (SE)*	**Chained MI (combining comorbidities)** *Coefficient (SE)*
Birth Weight	-0.00379*** (0.000133)	-0.00370*** (0.000126)	-0.00403*** (0.000138)	-0.00391*** (0.000130)	-0.00375*** (0.000133)	-0.00367*** (0.000126)
Year of Birth	0.0185 (0.0108)	0.00896 (0.0104)	0.0253* (0.0111)	0.0134 (0.0104)	0.0164 (0.0107)	0.00705 (0.0103)
Female	-0.868*** (0.0694)	-0.837*** (0.0676)	-0.896*** (0.0706)	-0.885*** (0.0678)	-0.877*** (0.0692)	-0.847*** (0.0674)
Institutional Delivery	0.716*** (0.102)	0.752*** (0.0993)	0.682*** (0.103)	0.718*** (0.0988)	0.719*** (0.102)	0.754*** (0.0991)
Twin	-0.104 (0.133)	-0.141 (0.128)	-0.140 (0.134)	-0.194 (0.127)	-0.0921 (0.133)	-0.130 (0.128)
High Risk on Day 1	-0.805*** (0.133)	-0.946*** (0.130)	-0.477*** (0.137)	-0.578*** (0.133)	-0.836*** (0.134)	-0.975*** (0.130)
Baby Feeding Problem	-1.128*** (0.239)	-1.190*** (0.231)	-1.197*** (0.241)	-1.232*** (0.230)	-1.040*** (0.185)	-1.095*** (0.179)
Mother Feeding Problem	-0.601** (0.215)	-0.522* (0.212)	-0.727*** (0.214)	-0.619** (0.212)	-0.571** (0.215)	-0.500* (0.212)
CHW Present At Delivery	0.287** (0.0999)	0.288** (0.0975)	0.192 (0.103)	0.254** (0.0971)	0.312** (0.1000)	0.311** (0.0976)
Delayed Breastfeeding	-1.269	-1.460*	-1.600*	-1.395*	-1.348	-1.526*

	(0.738)	(0.642)	(0.738)	(0.643)	(0.736)	(0.640)
Birth Asphyxia	0.0655 (0.153)	0.0745 (0.151)	0.0458 (0.152)	0.0370 (0.149)		
Congenital Anomaly	-1.014 (0.881)	-0.893 (0.854)	-1.070 (0.898)	-0.851 (0.862)		
Umbilical Sepsis	0.652 (0.672)	0.554 (0.660)	0.701 (0.675)	0.562 (0.658)		
Conjunctivitis	0.752 (0.474)	0.792 (0.472)	0.715 (0.456)	0.843 (0.474)		
Unexplained Fever	0.413* (0.200)	0.382 (0.196)	0.372 (0.200)	0.354 (0.195)		
Hemorrhage	-1.019 (0.601)	-0.995 (0.600)	-1.076 (0.610)	-1.070 (0.609)		
Hypothermia	-0.397 (0.350)	-0.393 (0.343)	-0.406 (0.353)	-0.377 (0.340)		
Neonatal Sepsis	0.110 (0.238)	0.133 (0.232)	0.192 (0.240)	0.196 (0.232)		
Pneumonia Only	0.386 (0.256)	0.337 (0.242)	0.410 (0.253)	0.355 (0.240)		
Bacterial Skin Infection	0.327 (0.257)	0.323 (0.252)	0.263 (0.250)	0.308 (0.251)		
Jaundice	0.406 (0.346)	0.388 (0.343)	0.249 (0.336)	0.372 (0.340)		
Bad Obstetric History	-0.475*** (0.111)	-0.444*** (0.111)	-0.468*** (0.113)	-0.408*** (0.111)		
Tobacco Use During Pregnancy	0.262** (0.0890)	0.280** (0.0886)	0.280** (0.0893)	0.283** (0.0883)	0.247** (0.0892)	0.266** (0.0888)
Number of CHW visits						
1-5 visits	-0.809*** (0.153)	-0.944*** (0.150)	-0.825*** (0.156)	-0.897*** (0.149)	-0.831*** (0.153)	-0.962*** (0.150)
5-10 visits	-0.317* (0.126)	-0.376** (0.123)	-0.189 (0.125)	-0.240 (0.123)	-0.310* (0.126)	-0.367** (0.124)
10+ visits	-0.390	-0.225	0.00292	0.139	-0.370	-0.207

	(0.209)	(0.202)	(0.211)	(0.203)	(0.211)	(0.204)
Mother's Education						
1-4/Literate	0.204	0.184	0.259	0.222	0.215	0.194
	(0.175)	(0.174)	(0.176)	(0.173)	(0.176)	(0.175)
5-10 class	0.436***	0.429***	0.437***	0.436***	0.443***	0.436***
	(0.128)	(0.127)	(0.128)	(0.127)	(0.128)	(0.127)
>= 11 class	0.922***	0.919***	0.965***	0.935***	0.935***	0.931***
	(0.148)	(0.148)	(0.148)	(0.148)	(0.149)	(0.148)
Father's Education						
1-4/Literate	0.0331	0.0180	0.0468	0.0314	0.0649	0.0492
	(0.153)	(0.154)	(0.154)	(0.153)	(0.153)	(0.153)
5-10 class	0.321*	0.322*	0.343**	0.352**	0.336**	0.336**
	(0.128)	(0.128)	(0.129)	(0.128)	(0.128)	(0.128)
>= 11 class	0.494***	0.500***	0.567***	0.551***	0.512***	0.517***
	(0.140)	(0.140)	(0.141)	(0.139)	(0.140)	(0.140)
Parity						
1 preg.	0.552***	0.610***	0.577***	0.609***	0.481***	0.542***
	(0.0776)	(0.0760)	(0.0772)	(0.0757)	(0.0765)	(0.0749)
2 preg.	0.480***	0.517***	0.540***	0.513***	0.310**	0.355**
	(0.116)	(0.115)	(0.116)	(0.115)	(0.112)	(0.111)
3 preg.	0.122	0.143	0.239	0.179	-0.130	-0.0976
	(0.190)	(0.189)	(0.190)	(0.189)	(0.188)	(0.187)
>=4 preg.	0.261	0.222	0.318	0.219	-0.0467	-0.0695
	(0.329)	(0.327)	(0.338)	(0.327)	(0.326)	(0.323)
Gestational Age (weeks)			0.223***	0.238***		
			(0.0230)	(0.0226)		
Number of Comorbidities					0.00838	0.0158
					(0.0779)	(0.0767)
Intercept	-19.14	-0.114	-41.07	-18.06	-14.90	3.640
	(21.69)	(20.87)	(22.36)	(20.85)	(21.39)	(20.64)
adj. R2	0.126	.	0.138	.	0.124	.
N	15422	16214	14970	16214	15422	16214

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 7(c)

Table 7(c). Multiple Linear Regression of Exponential Growth Velocity (g/kg/day) on Predictors						
Predictor	**Complete Case Analysis** ***Coefficient (SE)***	**Chained MI** ***Coefficient (SE)***	**Complete Case Analysis (controlling for Gestational Age)** ***Coefficient (SE)***	**Chained MI (controlling for Gestational Age)** ***Coefficient (SE)***	**Complete Case Analysis (combining comorbidities)** ***Coefficient (SE)***	**Chained MI (combining comorbidities)** ***Coefficient (SE)***
Birth Weight	-0.00291*** -0.0000955	-0.00285*** -0.000091	-0.00312*** -0.0000895	-0.00301*** -0.0000917	-0.00288*** -0.0000953	-0.00282*** -0.0000907
Year of Birth	0.015 -0.00881	0.00763 -0.00844	0.0222* -0.00879	0.0112 -0.0084	0.0138 -0.00873	0.00651 -0.00837
Female	-0.669*** -0.0545	-0.644*** -0.0532	-0.685*** -0.0545	-0.682*** -0.0532	-0.676*** -0.0544	-0.652*** -0.0531
Institutional Delivery	0.578*** -0.0817	0.607*** -0.0794	0.549*** -0.082	0.580*** -0.079	0.583*** -0.0815	0.611*** -0.0793
Twin	-0.0678 -0.106	-0.095 -0.102	-0.0855 -0.107	-0.137 -0.102	-0.0603 -0.106	-0.0886 -0.102
High Risk on Day 1	-0.680*** -0.108	-0.792*** -0.105	-0.419*** -0.11	-0.500*** -0.107	-0.714*** -0.108	-0.823*** -0.105
Baby Feeding Problem	-0.943*** -0.195	-0.992*** -0.188	-1.019*** -0.195	-1.025*** -0.188	-0.879*** -0.148	-0.922*** -0.143
Mother Feeding Problem	-0.520** -0.175	-0.460** -0.173	-0.619*** -0.175	-0.537** -0.172	-0.501** -0.177	-0.447* -0.174
CHW Present At Delivery	0.254** -0.0814	0.253** -0.0793	0.168* -0.0818	0.226** -0.0789	0.274*** -0.081	0.272*** -0.0788
Delayed Breastfeeding	-1.134 -0.633	-1.261* -0.55	-1.402* -0.642	-1.210* -0.552	-1.206 -0.632	-1.321* -0.548
Birth Asphyxia	0.0179	0.0269	0.000884	-0.00287		

	-0.11	-0.109	-0.11	-0.108		
Congenital Anomaly	-1.098	-1	-1.149	-0.967		
	-0.704	-0.683	-0.721	-0.692		
Umbilical Sepsis	0.281	0.207	0.314	0.213		
	-0.349	-0.344	-0.35	-0.343		
Conjunctivitis	0.598	0.629	0.592	0.67		
	-0.377	-0.376	-0.37	-0.378		
Unexplained Fever	0.293	0.27	0.261	0.247		
	-0.169	-0.166	-0.169	-0.165		
Hemorrhage	-0.806	-0.789	-0.848	-0.848		
	-0.483	-0.482	-0.49	-0.49		
Hypothermia	-0.512*	-0.501*	-0.532*	-0.488*		
	-0.241	-0.236	-0.242	-0.234		
Neonatal Sepsis	0.055	0.0766	0.129	0.126		
	-0.192	-0.187	-0.194	-0.186		
Pneumonia Only	0.312	0.274	0.33	0.289		
	-0.201	-0.19	-0.201	-0.189		
Bacterial Skin Infection	0.273	0.268	0.233	0.256		
	-0.206	-0.203	-0.203	-0.202		
Jaundice	0.315	0.301	0.199	0.288		
	-0.279	-0.277	-0.273	-0.274		
Bad Obstetric History	-0.376***	-0.353***	-0.379***	-0.324***		
	-0.091	-0.0907	-0.091	-0.0906		
Tobacco Use During Pregnancy	0.221**	0.235***	0.234**	0.237***	0.212**	0.226**
	-0.0716	-0.0713	-0.0718	-0.0711	-0.0717	-0.0714
Number of CHW visits						
1-5 visits	-0.652***	-0.762***	-0.684***	-0.725***	-0.667***	-0.773***
	-0.125	-0.122	-0.125	-0.121	-0.125	-0.122
5-10 visits	-0.241*	-0.287**	-0.151	-0.179	-0.233*	-0.278**
	-0.0998	-0.0979	-0.0993	-0.098	-0.0998	-0.098
10+ visits	-0.350*	-0.22	-0.0139	0.0685	-0.328*	-0.2
	-0.16	-0.156	-0.157	-0.158	-0.161	-0.156
Mother's Education						

1-4/Literate	0.172	0.156	0.216	0.186	0.182	0.166
	-0.144	-0.143	-0.144	-0.142	-0.145	-0.143
5-10 class	0.343***	0.339***	0.344***	0.344***	0.349***	0.345***
	-0.102	-0.101	-0.102	-0.101	-0.102	-0.101
>= 11 class	0.734***	0.732***	0.779***	0.745***	0.744***	0.742***
	-0.119	-0.119	-0.119	-0.119	-0.12	-0.119
Father's Education						
1-4/Literate	0.0271	0.0154	0.0339	0.0261	0.0565	0.0441
	-0.125	-0.125	-0.126	-0.125	-0.125	-0.125
5-10 class	0.247*	0.247*	0.260*	0.271**	0.261*	0.261*
	-0.104	-0.104	-0.104	-0.103	-0.104	-0.104
>= 11 class	0.368**	0.374**	0.433***	0.415***	0.385***	0.391***
	-0.114	-0.114	-0.115	-0.114	-0.115	-0.114
Parity						
1 preg.	0.434***	0.480***	0.470***	0.480***	0.382***	0.430***
	-0.0664	-0.0654	-0.0618	-0.0652	-0.065	-0.0639
2 preg.	0.381***	0.411***	0.441***	0.408***	0.260**	0.296**
	-0.0956	-0.0949	-0.0948	-0.0949	-0.0916	-0.0908
3 preg.	0.0985	0.116	0.207	0.144	-0.0853	-0.0594
	-0.158	-0.157	-0.156	-0.156	-0.155	-0.154
>=4 preg.	0.163	0.129	0.218	0.127	-0.0584	-0.0803
	-0.273	-0.271	-0.28	-0.271	-0.268	-0.266
Gestational Age (weeks)			0.175***	0.189***		
			-0.0179	-0.0181		
Number of Comorbidities					-0.0338	-0.0267
					-0.0522	-0.0514
Intercept	-15.33	-0.657	-36.16*	-14.88	-12.88	1.521
	-17.63	-16.91	-17.65	-16.85	-17.47	-16.76
adj. R2	0.118	.	0.136	.	0.116	.
N	15422	16214	14970	16214	15422	16214

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Regression Diagnosis

Standard checks of the regression were carried out. Results are presented here only for the simplest complete case analysis model, for absolute change in weight as an outcome.

Heteroskedasticity was checked with White's test. The null hypothesis of homoskedasticity was rejected, so Huber-White robust standard errors were used as mentioned in the previous section.

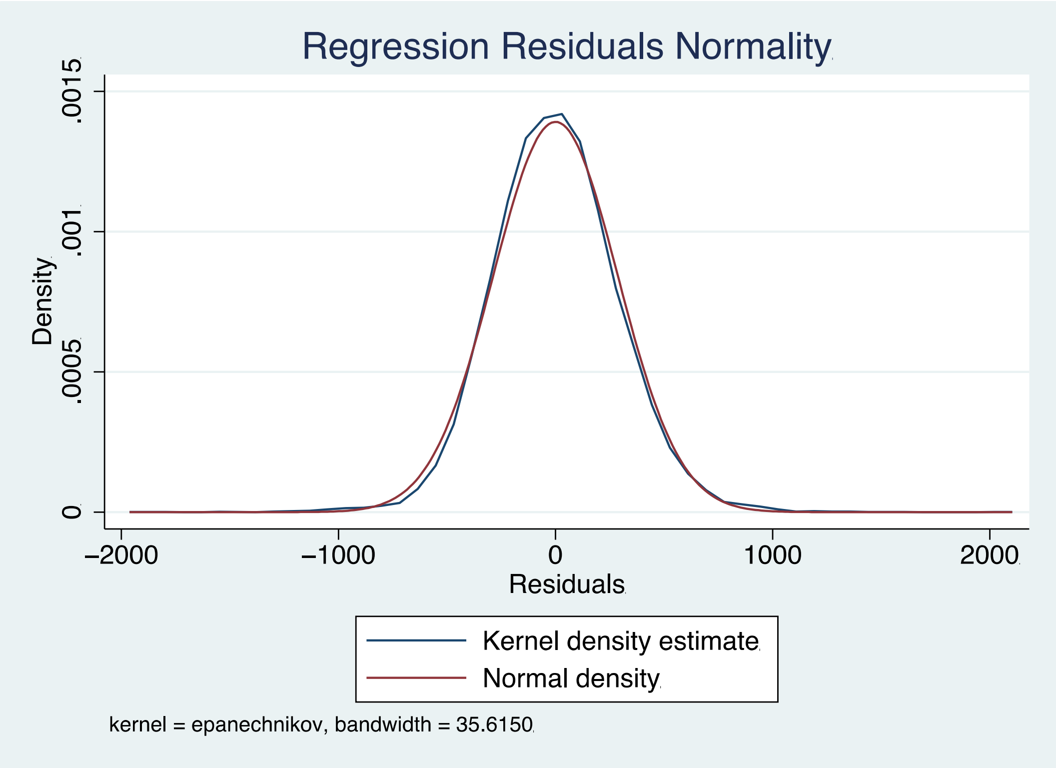
The regression was checked for multicollinearity, which is a high correlation between predictor variables. Multicollinearity can cause the model to be highly sensitive to inclusion or exclusion of variables, as well as weaken overall statistical power. It can be checked by examining the Variance Inflation Factors of predictor variables, which measure correlation between predictors. VIFs under 5 are generally considered acceptable. As shown in Table 8, VIFs for all predictors are below 5, with a mean VIF of just 1.8.

Table 8. Regression Diagnosis for Multicollinearity

Variable	Variance Inflation Factor
Year of Birth	4.02
Father Education >=11	3.87
Mother's Education >= 11	3.79
Father Education=5-10	3.63
Mother's Education 5-10	3.38
CHW visits 1-5	3.37
CHW visits 5-10	3.24
Twins	2.68
Institutional delivery	2.23
Baby Feeding Problem	2.15
High Risk on Day 1	2.15
Neonatal Sepsis	2.09
CHW Present at Delivery	1.9
Father's Education=Literate/1-4	1.69
CHW visits >=10	1.66
Parity = 2	1.4
Conjunctivitis	1.33
Jaundice	1.32
Mother's Education=Literate/1-4	1.32
B irth Weight	1.3
Parity = 1	1.29
Bad Obstetric History	1.24

Parity = 3	1.22
Parity = 4	1.14
Hypothermia	1.09
Tobacco Use in Pregnancy	1.08
Birth Asphyxia	1.05
Mother Feeding Problem	1.05
Bacterial Skin Infection	1.04
Umbilical Sepsis	1.03
Unexplained Fever	1.02
Sex	1.01
Delayed Breastfeeding	1.01
Hemorrhage	1.01
Pneumonia	1.01
Congenital Anomaly	1
Mean VIF	**1.83**

Another assumption to check of multiple regression is that the residuals are normally distributed. Again for the simplest model with complete case analysis, the residuals are shown below in comparison to a normal density. There is very close agreement, showing that this assumption is likely not violated much.



Remaining Open Questions for Analysis

1) Considering that the data recorded is panel data, with observations on day 1, 7, 15, 21, and 28, instead of conducting the analysis with ordinary linear regressions, we could

use a Generalised Estimating Equation (GEE) or a linear mixed-model. However, interpreting the results of such analysis seems somewhat more challenging and perhaps less physically meaningful.

2) The choice of which outcome variable to use for the final analysis is open. The use of absolute weight change is easier to interpret, but it may be more dependent on the background socioeconomic and genetic conditions of the target population, and so may yield less reproducible/generalisable conclusions for the wider community. If the goal is to present observations about weight change in this community, then weight change may be appropriate. However, if the goal is to generally determine the factors affecting weight gain, then weight gain velocity may be a more appropriate metric, since it is "standardized" as a fraction of neonate weight.
3) The choice of which model(s) to include is also open.
    a. After reviewing the literature, it is not exactly clear to me whether or not gestational age would be a collider variable in our case.
    b. Also, given that most comorbidities are insignificant predictors of either weight gain or velocity, combining them into one variable to simplify the model seems reasonable. However, some past literature has indicated that infections and similar comorbidities *do* have a significant weight gain effect, so leaving comorbidities disaggregated may help contradict that literature.
    c. Finally, all models have been presented with Complete Case Analysis and Chained MI results, to show that both ways of dealing with missing data do not vary too much in the end. However, presenting both results seems excessive.