Machine Learning Model Demonstrates Stunting at Birth and Systemic Inflammatory Biomarkers as Predictors of Subsequent Infant Growth – A Four-year Prospective Study.

Elizabeth Harrison
University of Virginia
Sana Syed (ss8xj@virginia.edu)
University of Virginia
Lubaina Ehsan
University of Virginia
Najeeha T. Iqbal
Aga Khan University
Kamran Sadiq
Aga Khan University
Fayyaz Umrani
Aga Khan University
Sheraz Ahmed
Aga Khan University
Najeeb Rahman
Aga Khan University
Sadaf Jakhro
Aga Khan University
Z. Jennie
university of virginia
Molly Hughes
university of virginia
S. Asad Ali
Aga Khan University

Research article

Keywords: Infant Growth, Systemic Inflammatory Biomarkers, Growth Predictors

Posted Date: June 26th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-37878/v1

Version of Record: A version of this preprint was published on October 30th, 2020. See the published version at https://doi.org/10.1186/s12887-020-02392-3.

Abstract

Background: Stunting affects up to one-third of children in low-to-middle income countries (LMICs) and has been correlated with cognitive decline and vaccine immunogenicity. Infants with stunting often have growth refractory to nutritional interventions. Early identification of at-risk infants is critical for early intervention and prevention of subsequent morbidity. The aim of this study was to investigate patterns of growth in infants up through 48 months of age, as well as potential predictors of growth.

Methods: Height-for-age z-scores (HAZ) of children from Matiari (rural site in Pakistan) at birth, 18 months, and 48 months were obtained. Results of serum-based biomarkers collected at 6 and 9 months were recorded. A descriptive analysis of the population was followed by biomarker prediction assessment via machine learning methods; random forest classification bins of z-scores and random forest regression for continuous z-scores.

Results: Of 107 children who were followed up to 48 months, about half were stunted (HAZ < -2) at birth which increased to 54% by 48 months of age. HAZ at 18 months had a positive linear correlation with HAZ at 48 months, while HAZ at birth was less strongly correlated with HAZ at either 18 or 48 months. Stunting status for the majority of children at 48 months was found to be the same as at 18 months. Most children with large gains started off stunted or severely stunted, while all of those with notably large losses were not stunted at birth. Random forest models identified HAZ at birth as the most important feature in predicting HAZ at 18 months. Of the biomarkers, AGP (Alpha- 1-acid Glycoprotein), CRP (C-Reactive Protein), and IL1 (interleukin-1) were assigned the highest importance rankings across both the classification and regressor models.

Conclusion: We demonstrated that HAZ at birth is a strong predictor of subsequent growth in infants up through 48 months of age. Biomarkers of systemic inflammation, AGP, CRP, IL1, were also strong predictors of growth outcomes. These findings provide support for continued focus on interventions at birth and early infancy in children at risk for stunting who live in resource-constrained regions of the world.

Background

Stunting affects up to one-third of children in low-to-middle income countries (LMICs) (1). It is indicative of a failure to achieve one's genetic potential for height (more than two standard deviations below the World Health Organization international standards for growth) (2, 3). Long-term devastating consequences of stunting have been reported, including permanent cognitive impairments, oral vaccine failure, and diminished immunocompetence (1, 4). It further accounts for 1.2 million deaths per year among children under 5 years of age (1). Global income has been estimated to increase by $176.8 billion per year if linear growth failure is eliminated (5). Linear growth improvement has been reported by previous studies to be refractory to nutritional interventions (6, 7). This prompts the need to explore those

at-risk individuals in order to eradicate factors leading to stunting, and warrants well-designed trials to elucidate any and all food-based interventions that might have growth-promoting potential (6).

The age of stunting has direct implications for the progression of growth, as well as the timing and nature of appropriate interventions. Intrauterine growth restriction and small size at birth are strongly associated with risk of stunting at 24 months of age (8). Most relevant studies have shown that major linear growth failure occurs in the first 48 months of life; beyond this age, catch-up growth is rare due to a lack of change in nutrition and environment for older children (9, 10). A large body of evidence suggests that the first 1000 days from conception is a critical window in which interventions to address malnutrition will be most effective; however, little is known about the impact on linear growth of nutritional interventions in children greater than 2 years of age (6).

The early identification of at-risk infants is critical for early intervention and prevention of subsequent morbidity. Previous studies have shown increased concentrations of inflammatory biomarkers and decreased concentrations of anabolic growth factors such as insulin-like growth factor-1 (IGF-1) to be associated with stunting (4). Such studies often utilize parametric methods in their data analyses, even though nonparametric approaches - such as random forests - frequently outperform parametric approaches in studies with a larger number of variables than observations (11).

The aim of our study was to investigate patterns of growth in infants up through 48 months of age, as well as potential predictors of growth such as systemic biomarkers and anthropometric measurements taken at birth.

Methods

Data Collection

As part of a 4-year prospective study, researchers at Aga Khan University in Pakistan collected data for 380 children from the rural village of Matiari, Sindh, Pakistan (12). The data comprised a combination of demographic, biometric, and laboratory measures. Variables included monthly anthropometric measurements collected from birth through the twentieth month of life. For this study, consent was also obtained for additional weight and height measurements to be taken at 48 months of age. All anthropometric measurements were converted into z-scores using the World Health Organization child growth standards.

Ethics Approval

This study was approved the Ethical Review Committee of Aga Khan University in Karachi, Pakistan (informed consent obtained from parents and/ or guardians).

Analysis Overview

Our aims were (1) to characterize this population descriptively at the 48-month follow-up data for this population and (2) to better identify – via random forest analysis – those anthropometric measurements and biomarker levels collected in the first year of life that might serve as predictors of subsequent growth. Visualizations used to aid descriptive analysis included scatter plots, bar plots, and spaghetti plots. Random forests analysis identified predictors of growth via classification and regression. With the aid of additional visualizations, these results were also used to rank the predictors of linear growth at 20 months of life. The predictability of the top 25 variables was then estimated using a linear model. Data preparation, modeling, and analysis were all completed using the Python coding language in Jupyter Notebook, an open-source development environment.

Descriptive Analysis

Data exploration initially focused on 48 month outcomes. Of the original cohort (n = 380), which was followed for 24 months, 112 infants also participated in the follow-up study up till 48 months of age. 107 infants had sufficient anthropometric data to be included in analysis. Stunting and growth failure in this study were evaluated with height-for-age z-scores (HAZ). Stunting was defined as HAZ two standard deviations below the mean (HAZ < -2).

Mean HAZ was calculated across the follow-up population at three time points: (1) at enrollment (< 1 month of age), (2) at 18 months of age, and (3) at 48 months of age. Patterns in the distribution of stunting across both sexes were examined at each point. Then, the influence of location was examined, including any patterns associated with village or Union Council. Subsequently, based on the same three time points, the population was divided into subgroups based on stunting status at each of these study visits. This allowed for an examination of general growth trends using a categorical variable. Growth trends were also evaluated using linear regression plots and correlation coefficients, with x and y values based on individual children's raw HAZ at each of the three time points.

In order to assess the change in growth, this study further examined growth trends based on the changes in HAZ over time (delta HAZ). Delta HAZ were calculated by subtracting the z-scores at 18 and 48 months from those given at the first clinic visit. With a slightly smaller subset of 101 children, a spaghetti plot was used to identify growth patterns in the follow-up cohort using monthly HAZ measurements from the first 18 months of life, as well as from the 48-month follow-up visit. Relevant delta HAZ outliers were then highlighted using different colors based on whether their delta HAZ was notably positive or negative.

Random Forest Models

This study's final models were designed to be interpretable, with a significantly reduced set of predictors. All models were created using sklearn's Random Forest packages. List of biomarkers and cytokines included in the random forest model is provided in Additional File 1.

Identification of variables highly predictive of stunting was approached in two ways: 1) with a classification model using stunted versus not-stunted as outcomes; and 2) regression using HAZ at 18 months as the outcome variable. For both approaches, an 80 – 20 test-train split was used. To minimize

bias, children who participated in the follow-up were divided randomly across training and testing groups. All numeric variables were scaled using sklearn's min_max_scaler.fit_transform( ) function.

In classification, random forest hyperparameters were optimized using a grid search. This grid search included "n" estimators ranging from 5 to 300, max features ranging from 25 to 106 (all features), max depths ranging from 5 to "None," minimum sample splits of 2 and 4, and minimum sample leaves of 1 and 2. The grid search comprised 300 iterations with 5-fold cross validation. Optimized parameters included max depth at 100, max features at 106 (all features), "n" estimators at 80, minimum sample leaf number at 2, and minimum sample split at 4. All other hyperparameters were set to the function's default. Feature importance results were then extracted and the top 35 features of the forest were plotted using a labeled bar chart with lines over each bar representing the inter-tree variability of each feature.

In regression, performance of a baseline model using all default hyperparameters was compared to that of a model using hyperparameters optimized with a similar grid search to that noted above. Performance was relatively comparable and highly dependent on random state, so the baseline random forest regressor model was used for feature analysis. Again, the top 35 features of importance were extracted and plotted. These were then compared to the classification model's outputs.

Results

Descriptive Analysis

A total of 107 children who were followed up to 48 months were included from the follow-up cohort during initial descriptive analysis with 46% of these being male (Tables 1 and 2). Of this cohort, 51% were found to have been stunted at birth with 51% identified as male. At 18 months of age, the percentage of stunted children rose to 64%, but the male percentage of this subgroup decreased to 46%. By the 48-month visit, the percentage of stunted children had dropped back down to 54%, with only 40% of these being male. Using scatter plots and linear fit regression lines, HAZ at birth, 18 months, and 48 months were compared (Fig. 1). While HAZ at birth were weakly linearly correlated with HAZ at either 18 months or 48 months ($r = 0.376$; $p = 0.0001$ and $r = 0.162$; $p = 0.0954$), respectively). HAZ at 18 months showed a stronger positive linear correlation with HAZ at 48 months ($r = 0.604$; $p < 0.0001$).

Table 1
Patient and Maternal Characteristics of the 107 infants followed till 48 months of age.

Characteristics	Frequency (%)
Gender, Male (%)	46
Preterm Birth (%)	52*
Advanced Maternal Age (%)	64*
Breastfed soon after birth (%)	98*
Literate Mother (%)	13*
* – values missing for 6 children	

Table 2
Anthropometric measurement based WHZ, HAZ and WAZ of the 107 infants followed till 48 months of age.

Anthropometric Measurements	Birth (mean, ± SD)	18 months of age (mean, ± SD)	48 months of age (mean, ± SD)
WHZ	-0.25, ± 1.30	-0.81, ± 0.92	-0.66, ± 0.89
HAZ	-2.09, ± 1.35	-2.37, ± 0.87	-2.11, ± 0.87
WAZ	-1.68, ± 1.17	-1.73, ± 0.91	-1.73, ± 0.83
Key: WHZ – Weight for Height z-score; HAZ – Height for Age z-score; WAZ – Weight for Age z-score; SD – Standard Deviation			

The 107 children were examined by dividing them into groups based on their stunting at these same time points. As shown in Table 3, of the 107 children, 20 children had HAZ above − 2 at all three time points, while 30 children who began stunted (i.e. had HAZ less than − 2) remained stunted at their 18 and 48 follow ups. Of the 55 children who were stunted at birth, 12 grew out of being stunted by 18 months (and remained in the normal HAZ range), and an additional 11 children grew out of being stunted 48 months. Of the 52 children who were not stunted at birth, 21 fell into the stunted range by 18 months and remained stunted at their follow-up 2 years later, and an additional 5 children fell into the stunted range between their 18-month and 48-month checks.

Table 3
Groups based on stunting status and follow up time points of the 107 infants followed till 48 months of age.

Birth	18 months of age (mean, ± SD)	48 months of age (mean, ± SD)	Number of Children (total n = 107)
Not Stunted	Not Stunted	Not Stunted	20
		Stunted	5
	Stunted	Not Stunted	6
		Stunted	21
Stunted	Not Stunted	Not Stunted	12
		Stunted	2
	Stunted	Not Stunted	11
		Stunted	30
Key: SD – Standard Deviation			
Figure Titles			

Using a more granular approach, 101 individual children's growth trends were compared using a spaghetti plot (Fig. 2). Six children from the above set were excluded from this analysis since they had less frequent visits in the first 2 years of the study. As shown in the plot, there were 9 children whose delta HAZ were over 2 (z-scores increased by at least 2 points) and 8 children whose delta HAZ were under − 2 (z-scores decreased by at least 2 points). All those with a positive change in z-score of 2 or more were noted to be stunted at their first study visit. Conversely, all those with a loss in their HAZ of 2 or more were notably not stunted at their first study visit. While almost all of the children in this subgroup who dropped in their HAZ experienced their greatest losses in the first 2 years, most of the children in this subgroup who grew well experienced their greatest gains between their 18-month study visit and the follow-up 48-month visit. Across the entire follow-up cohort, most children appear to gain a little between the 18-month and 48-month visits, though generally remaining around the same z-score.

Random Forest Models

After creating a model with sklearn's Random Forest Classifier and optimizing hyperparameters using a grid search with cross-validation, relative feature importance was extracted and plotted for the top 35 features (Fig. 3). As shown, the only feature identified with high importance and an inter-tree variability line that does not cross zero was the raw HAZ calculated from anthropometry done at birth. Other relatively important features in this forest included AGP (Alpha- 1-acid Glycoprotein) and CRP (C-Reactive Protein) biomarker levels at 9 months, as well as the lL1 (interleukin-1) biomarker; however, all of these demonstrated significant variability, with some trees assigning these features low to zero importance.

With optimization, this random forest model predicted stunting status at 18 months of age in the set-aside testing set with 78% accuracy.

Using a similar approach with sklearn's Random Forest Regressor, 35 features were once again identified and plotted (Fig. 4). Although there are a few features other than HAZ in this model with variability lines that do not cross zero, these results were inconsistent across different runs during model development, all due to changes in the random state assigned to each regressor. However, it is notable that the levels of AGP at 9 months of age presents as a strong feature in results of the classification and the regression model, in addition to having this reduced variability between trees of the regressor's random forest. Again, CRP is identified as an important feature, though this time from 6-month measurements, as well as IL1.

A baseline random forest regression model using all default hyperparameters and a model using hyperparameters identified with a grid search performed comparably on an unseen testing set. Both models predicted HAZ at 18 months with a mean squared error between 0.7 and 0.8, depending on random state assignments.

Conspicuously, none of the models identified sex, gestational age, or any of the maternal factors as highly important features; the only one of these features to be included in the top 35 of either model was maternal literacy (in the random forest classifier model), and it was of minimal importance across all trees in the forest.

Discussion

The present study investigated the growth of infants up through 48 months of age, as well as the potential systemic biomarker and anthropometric predictors of growth among this population from a rural village at an LMIC. A descriptive analysis was followed by the utilization of nonparametric machine learning models, specifically a random forest analysis of growth predictors. The major results of this work include the following: (1) 51% of the infants were found to be stunted at birth; (2) a stronger correlation exists between HAZ at 18 and 48 months when compared to the correlation between HAZ at birth and either 18 or 48 months; (3) of all the systemic biomarkers and anthropometric measurements, HAZ at birth, AGP, CRP, and IL1 were found to be the strongest predictors of stunting.

Most studies have shown that major linear growth failure occurs in the first 48 months of life; beyond this age, catch-up growth is rare (9, 10). A large body of evidence suggests that the first 1000 days from conception is a critical window in which interventions to address malnutrition will be most effective, but little is known about the impact on linear growth of nutritional interventions in toddlers over the age of 2 years (6). In our study, HAZ at 18 months had a positive linear correlation with HAZ at 48 months, while HAZ at birth was less strongly correlated with HAZ at either 18 or 48 months of age. The positive correlation of HAZ at 18 and 48 months of age is supported by earlier reporting of minimal levels of catch-up growth from age 2 to 5 years (9). An increase in HAZ of children with stunting may also be a result of regression to the mean, as shown in previous studies (10).

HAZ at birth was shown to be a significantly strong predictor of growth followed by CRP, AGP, and IL1. Much emphasis has been placed on this first result from previous studies (i.e. early stunting as predictive of later stunting and neurodevelopmental outcomes) (13, 14). Such findings highlight the importance of interventions in the early months and years of life to prevent subsequent stunting and its consequences (14). It is important to note that nutritional interventions have shown little to no effect with regard to addressing neurodevelopmental outcomes and other long-term consequences if the child's stunting status is not also addressed (6). In order to design and construct trials that truly alleviate or reduce the consequences of stunting, the underlying factors that contribute to stunting itself need to be understood.

The serological biomarkers CRP, AGP, and IL1 succeeded stunting at birth in predicting growth among this study's cohort of children. CRP and AGP are acute phase proteins stimulated by the release of cytokines such as IL1, IL6, and TNF-α (Tumor necrosis factor-α) (15). CRP rises and declines rapidly during an acute phase response, whereas AGP rises more slowly (more than 24 hours after onset of inflammation) and remains elevated for longer (15–17). These findings are similar to a previous study by our group. In our previous work, we found significant correlations between flagellin- and lipopolysaccharide-specific (LPS-specific) IgA; serum CRP, AGP, and Reg1 (Regenerating Gene 1β) at 6 months; and MPO (myeloperoxidase) at 9 months. In the previous study, we found that higher anti-LPS IgA levels predicted greater declines in HAZ over the subsequent 18 months of follow-up (18). In contrast to this prior work, the current work utilizes a machine-learning model to investigate biomarkers as predictors of growth among infants.

Mixed association between inflammation and growth outcomes has also been reported previously. A study done among Zimbabwean infants (an LMIC setting) showed that levels of inflammatory biomarkers (CRP and AGP) measured at 6 weeks, 6 months, 12 months, and 18 months were consistently higher in children with stunting (HAZ < -2) versus healthy controls (defined as HAZ > -0.5) at 18 months (19). Further, among apparently healthy Zimbabwean infants with increased inflammatory biomarkers, the levels of anabolic hormone IGF-1 were low. This finding highlighted the significance of even low-grade inflammation with regard to poor growth outcomes (19). All findings provide support for continued focus on interventions at birth and early infancy in children at risk for stunting who live in resource-constrained regions of the world.

Knowledge about the biomarkers predictive of stunting is not only important from the perspective of constructing effective interventional trials, but also with regard to understanding the underlying pathology of stunting. The health of pregnant mothers has been shown to effect the infant at birth (20). It has been reported that stunting begins in utero and continues for at least the first 2 years of postnatal life; the period from conception to a child's second birthday (the first 1000 days) has therefore been identified as the most critical window of opportunity for interventions (13). Higher levels of inflammatory markers among infants could be due to ongoing inflammation in the pregnant mothers or to infections contracted during early life in the setting of poor sanitation and hygiene.

Several strengths of our study merit mention. Since this is a prospective study, all patients underwent biomarker collection and anthropometric measurements within a similar time frame. We followed a cohort of children for not only up through 24 months, but 48 months of age to assess growth outcomes and predictors and to answer the important questions regarding growth patterns between 2 and 5 years of age. The repeated measurements of length and height allowed our analysis evaluate growth status at birth, 18 months, and 48 months of age along with systemic biomarker levels, allowing us to assess the best predictors of growth beyond 2 years of age. Finally, we utilized a robust machine learning model to perform random forest analysis for the investigation of systemic biomarker and anthropometric growth predictors. Due to the prospective nature of this study, limitations included missing data points for Pearson correlations between known characteristics of participants who failed to return for follow-up or were missing biomarker results because of sample limitations. This also prompted the machine learning model to be designed based on sufficient data available at 18 months of age rather than 48 months of age.

An interesting question for future analysis will be the assessment of in-utero growth and inflammation as potential predictors of subsequent infant growth. Further, investigation of maternal factors – including systemic inflammation – might answer the important unanswered question regarding their role in the growth of the child along with stunting and its prevention.

Conclusion

We described the growth of infants up through 48 months of age and investigate the potential indicators of subsequent growth. While several of our findings (such as that HAZ at birth, AGP, CRP, and IL1 are predictive of subsequent growth) reiterate previous data, our work solidifies previous assessments of growth through 2 years of age and utilizes a robust a machine learning approach to confirm these measures as predictive of early infant growth patterns. This is significant, as it stresses the need to investigate maternal factors leading to stunting. It also highlights specific biomarkers that need to be factored in during construction of future trials targeted towards improvement of growth. These findings provide support for continued focus on interventions at birth and early infancy in children at risk for stunting who live in resource-constrained regions of the world.

Abbreviations

LMIC
low-to-middle income country
HAZ
Height-for-age z-score
AGP
Alpha- 1-acid Glycoprotein
CRP
C-Reactive Protein

IL1
interleukin-1
IGF-1
insulin-like growth factor-1
WHZ
Weight for Height z-score
HAZ
Height for Age z-score
WAZ
Weight for Age z-score
SD
Standard Deviation
GLP2
Glucagon-like peptide 2
HuEotaxin
Human Eotaxin
HuGCSF
Human Granulocyte-colony stimulating factor
HuIL4
Human Interleukin-8
HuIL7
Human Interleukin-7
HuIL8
Human Interleukin-8
HuIL9
Human Interleukin-9
HuILra
Human Interleukin-1 Receptor Antagonist
HuIP10
Human Interferon gamma-induced protein 10
HuPDGFb
Human Platelet Derived Growth Factor Subunit B
HuRANTES
Human RANTES (CCL5; C-C Motif Chemokine Ligand 5)
HuTNFa
Human Tumor necrosis factor-α
HuVEGF
Human Vascular Endothelial Growth Factor
LPSIgAOD
Lipopolysaccharide IgA Optical Density

LPSIgGOD
Lipopolysaccharide IgG Optical Density
MotherLiterate
Literacy Status of the Mother
MPO
Myeloperoxidase
NEO
Neopterin
Reg1Serum
Serum Regenerating Gene 1β
IL6
interleukin-6
TNF-α
Tumor necrosis factor-α

Declarations

Competing interests: None

Funding: This work was supported by the Bill and Melinda Gates Foundation grant OPP1066200 and OPP1066203 (SAA). The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Authors' Contributions

EHH, SS, SAA conceived and designed the study. EHH, SS developed and designed the methodology. EHH, NR, JZM analyzed the data. NTI, KS, FU, SA, SJ, MH contributed towards providing data and materials. EHH, LE wrote the paper. SS, SAA, NTI, JZM, SA, LE reviewed and edited the paper. SS, SAA supervised the overall study.

References

1. Black RE, Allen LH, Bhutta ZA, Caulfield LE, De Onis M, Ezzati M, et al. Maternal and child undernutrition: global and regional exposures and health consequences. The lancet. 2008;371(9608):243–60.
2. Organization WH. WHO child growth standards: length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: methods and development. 2006.
3. Organization WH. Nutrition. Stunting in a Nutshell [Available from: https://www.who.int/nutrition/healthygrowthproj_stunted_videos/en/.
4. Syed S, Manji K, McDonald C, Kisenge R, Aboud S, Sudfeld C, et al. Biomarkers of systemic inflammation and growth in early infancy are associated with stunting in young Tanzanian children.

Nutrients. 2018;10(9):1158.

5. Fink G, Peet E, Danaei G, Andrews K, McCoy DC, Sudfeld CR, et al. Schooling and wage income losses due to early-childhood growth faltering in developing countries: national, regional, and global estimates. Am J Clin Nutr. 2016;104(1):104–12.
6. Roberts JL, Stein AD. The Impact of Nutritional Interventions beyond the First 2 Years of Life on Linear Growth: A Systematic Review and Meta-Analysis. Advances in nutrition (Bethesda. Md). 2017;8(2):323–36.
7. Maleta K, Kuittinen J, Duggan MB, Briend A, Manary M, Wales J, et al. Supplementary feeding of underweight, stunted Malawian children with a ready-to-use food. J Pediatr Gastroenterol Nutr. 2004;38(2):152–8.
8. Thurow R. The First 1,000 Days: A Crucial Time for Mothers and Children-And the World. Breastfeed Med. 2016;11:416–8.
9. Desmond C, Casale D. Catch-up growth in stunted children: Definitions and predictors. PloS one. 2017;12(12):e0189135.
10. Leroy JL, Ruel M, Habicht JP, Frongillo EA. Using height-for-age differences (HAD) instead of height-for-age z-scores (HAZ) for the meaningful measurement of population-level catch-up in linear growth in children less than 5 years of age. BMC Pediatr. 2015;15:145.
11. Diaz-Uriarte R, Alvarez de Andres S. Gene selection and classification of microarray data using random forest. BMC Bioinform. 2006;7:3.
12. Iqbal NT, Sadiq K, Syed S, Akhund T, Umrani F, Ahmed S, et al. Promising biomarkers of environmental enteric dysfunction: a prospective cohort study in Pakistani children. Scientific reports. 2018;8(1):2966.
13. Prendergast AJ, Humphrey JH. The stunting syndrome in developing countries. Paediatrics international child health. 2014;34(4):250–65.
14. Prentice AM, Ward KA, Goldberg GR, Jarjou LM, Moore SE, Fulford AJ, et al. Critical windows for nutritional interventions against stunting. Am J Clin Nutr. 2013;97(5):911–8.
15. Jain S, Gautam V, Naseem S. Acute-phase proteins: As diagnostic tool. J Pharm Bioallied Sci. 2011;3(1):118–27.
16. Sproston NR, Ashworth JJ. Role of C-Reactive Protein at Sites of Inflammation and Infection. Front Immunol. 2018;9:754-.
17. Schultz DR, Arnold PI. Properties of four acute phase proteins: C-reactive protein, serum amyloid A protein, alpha 1-acid glycoprotein, and fibrinogen. Semin Arthritis Rheum. 1990;20(3):129–47.
18. Syed S, Iqbal NT, Sadiq K, Ma JZ, Akhund T, Xin W, et al. Serum anti-flagellin and anti-lipopolysaccharide immunoglobulins as predictors of linear growth faltering in Pakistani infants at risk for environmental enteric dysfunction. PloS one. 2018;13(3):e0193768-e.
19. Prendergast AJ, Rukobo S, Chasekwa B, Mutasa K, Ntozini R, Mbuya MN, et al. Stunting is characterized by chronic inflammation in Zimbabwean infants. PLoS One. 2014;9(2):e86928.

20. Neumann CG, Harrison GG. Onset and evolution of stunting in infants and children. Examples from the Human Nutrition Collaborative Research Support Program. Kenya and Egypt studies. Eur J Clin Nutr. 1994;48(Suppl 1):90–102.

21. Additonal Files

22. Additonal. **File 1: List of biomarkers and cytokines included in the random forest model**.

Figures

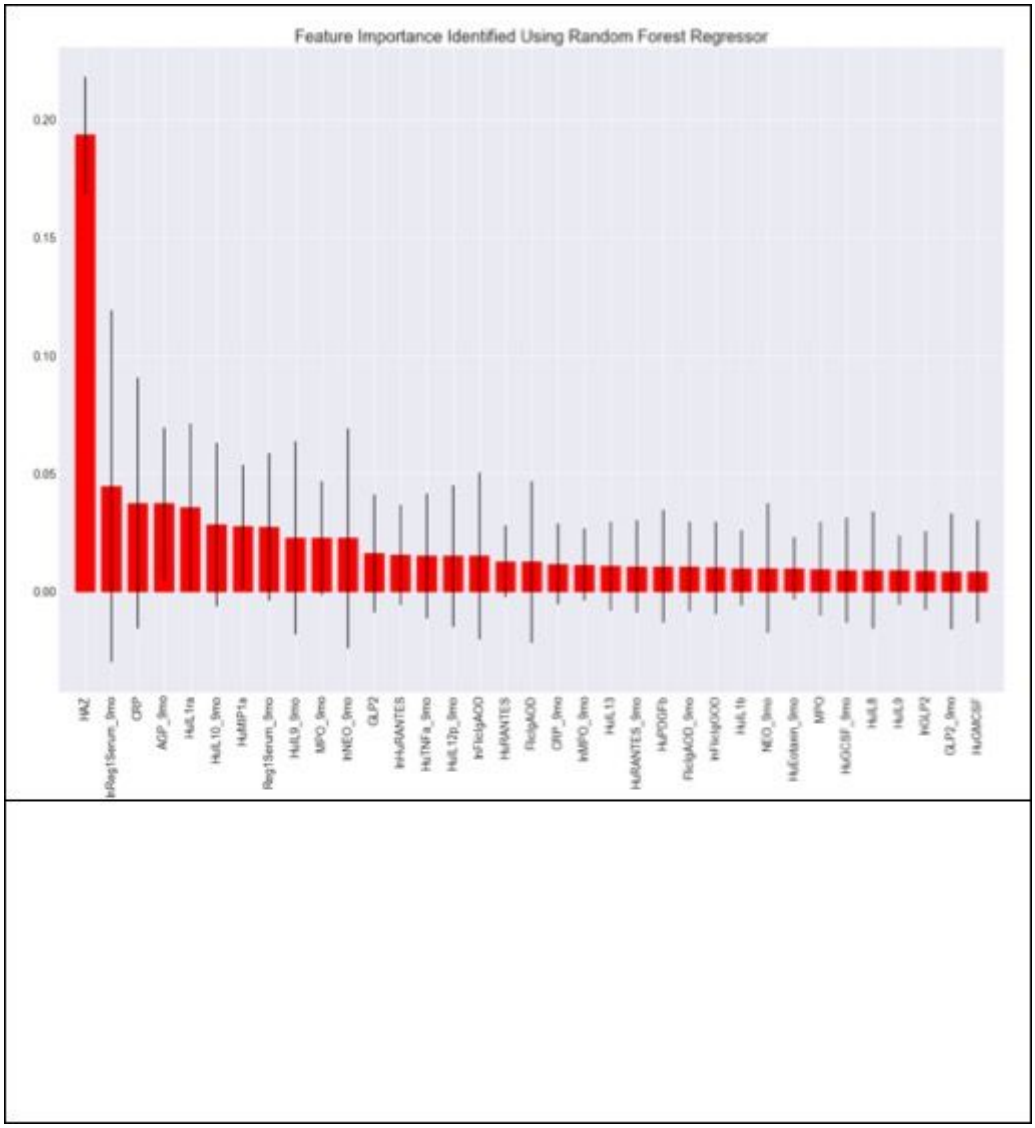

Figure 1

Random Forest Regressor based relative feature importance for the top 35 features predicting subsequent infant growth. Key (alphabetical): AGP – Alpha- 1-acid Glycoprotein; CRP – C-reactive Protein; GLP2 – Glucagon-like peptide 2; HAZ - Height for Age z-score; HuEotaxin – Human Eotaxin; HuGCSF – Human Granulocyte-colony stimulating factor; HuIL4 – Human Interleukin-8; HuIL7 – Human Interleukin-7; HuIL8 – Human Interleukin-8; HuIL9 – Human Interleukin-9; HuILra – Human Interleukin-1 Receptor

Antagonist; HuIP10 – Human Interferon gamma-induced protein 10;HuPDGFb – Human Platelet Derived Growth Factor Subunit B; HuRANTES – Human RANTES (CCL5; C-C Motif Chemokine Ligand 5); HuTNFa – Human Tumor necrosis factor-α; HuVEGF – Human Vascular Endothelial Growth Factor; LPSIgAOD – Lipopolysaccharide IgA Optical Density; LPSIgGOD – Lipopolysaccharide IgG Optical Density; MotherLiterate – Literacy Status of the Mother; MPO – Myeloperoxidase; NEO – Neopterin; Reg1Serum – Serum Regenerating Gene 1β. Note: 'ln' before a variable refers to natural logarithm, '_9mo' after a variable refers to the biomarker being collected at 9 months of follow-up, variables without mention of a time frame were collected at birth.

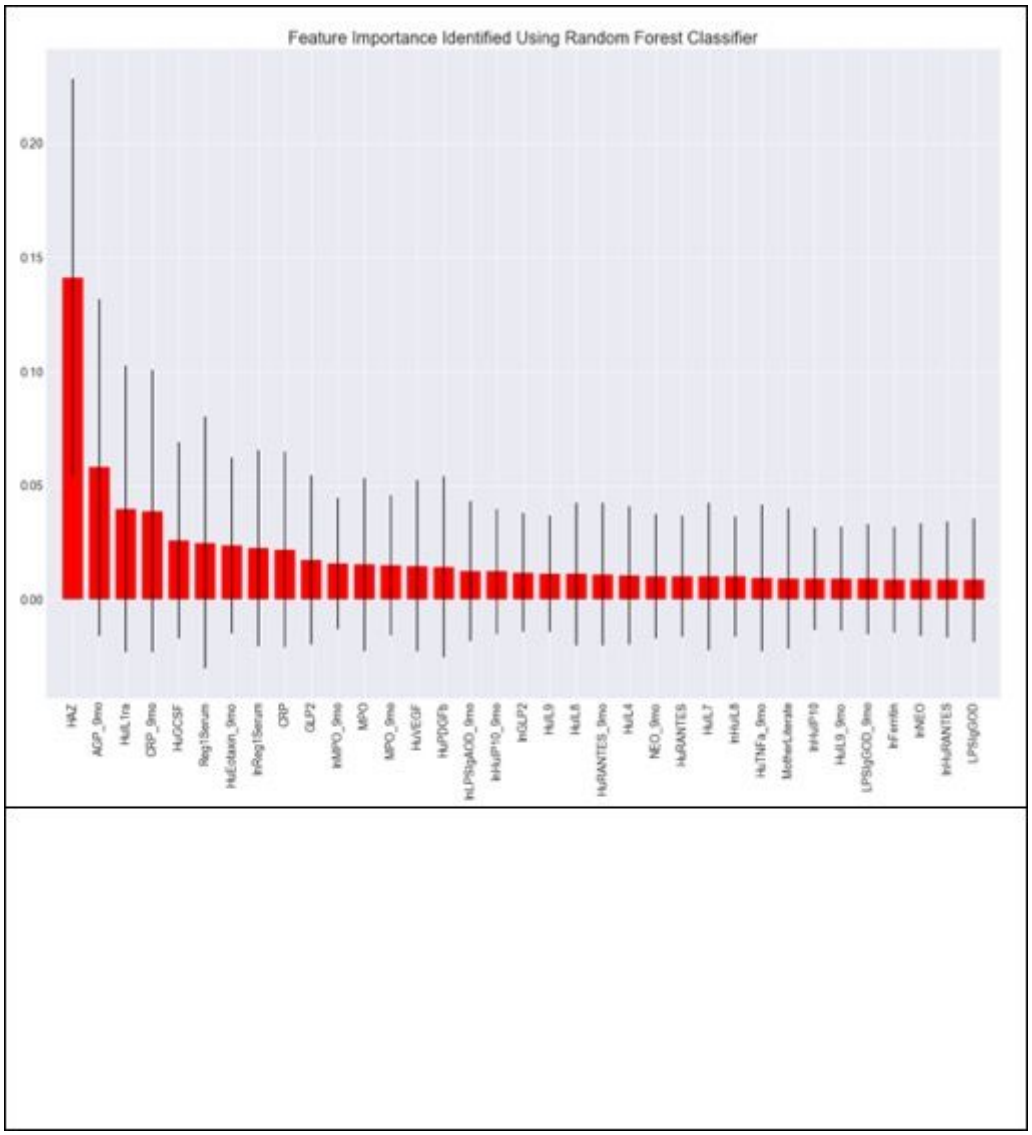

Figure 2

Random Forest Classifier based relative feature importance for the top 35 features predicting subsequent infant growth (hyperparameter optimization done via a grid search with cross-validation). Key (alphabetical): AGP – Alpha- 1-acid Glycoprotein; CRP – C-reactive Protein; GLP2 – Glucagon-like peptide 2; HAZ - Height for Age z-score; HuEotaxin – Human Eotaxin; HuGCSF – Human Granulocyte-colony stimulating factor; HuIL4 – Human Interleukin-8; HuIL7 – Human Interleukin-7; HuIL8 – Human Interleukin-8; HuIL9 – Human Interleukin-9; HuILra – Human Interleukin-1 Receptor Antagonist; HuIP10 –

Human Interferon gamma-induced protein 10;HuPDGFb – Human Platelet Derived Growth Factor Subunit B; HuRANTES – Human RANTES (CCL5; C-C Motif Chemokine Ligand 5); HuTNFa – Human Tumor necrosis factor-α; HuVEGF – Human Vascular Endothelial Growth Factor; LPSIgAOD – Lipopolysaccharide IgA Optical Density; LPSIgGOD – Lipopolysaccharide IgG Optical Density; MotherLiterate – Literacy Status of the Mother; MPO – Myeloperoxidase; NEO – Neopterin; Reg1Serum – Serum Regenerating Gene 1β. Note: 'ln' before a variable refers to natural logarithm, '_9mo' after a variable refers to the biomarker being collected at 9 months of follow-up, variables without mention of a time frame were collected at birth.

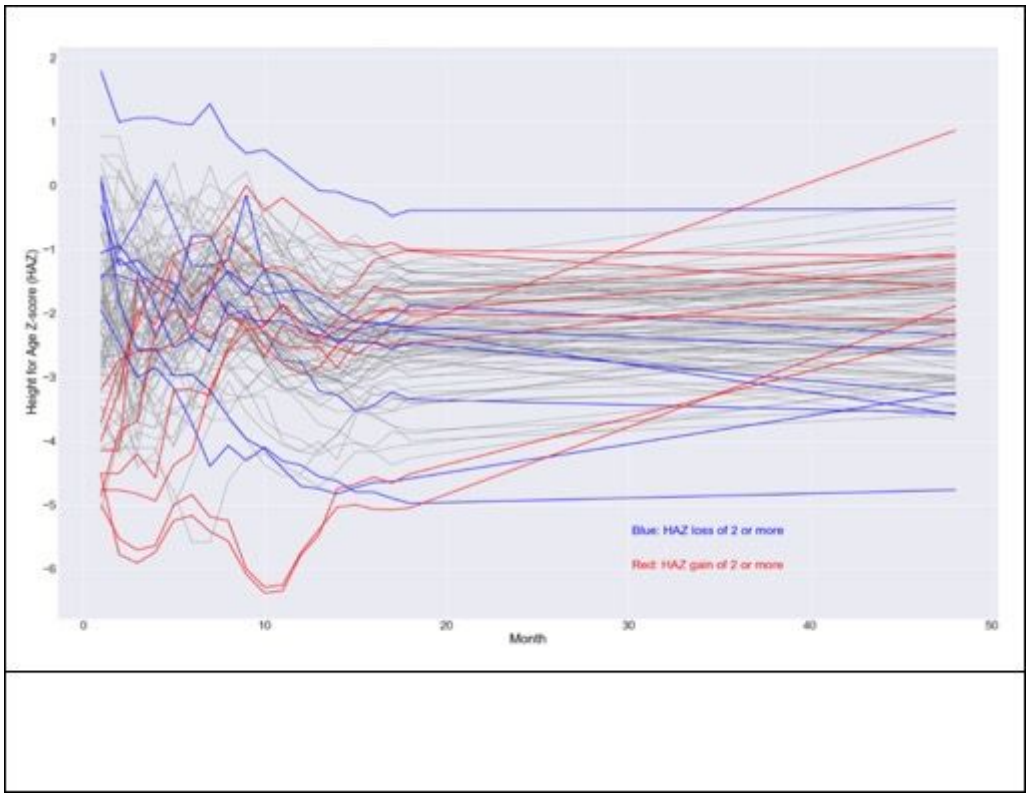

Figure 3

Spaghetti plot for growth trends of infants with follow up 48 months of age (6 patients excluded due to insufficient anthropometric data). Key: HAZ - Height for Age z-score.

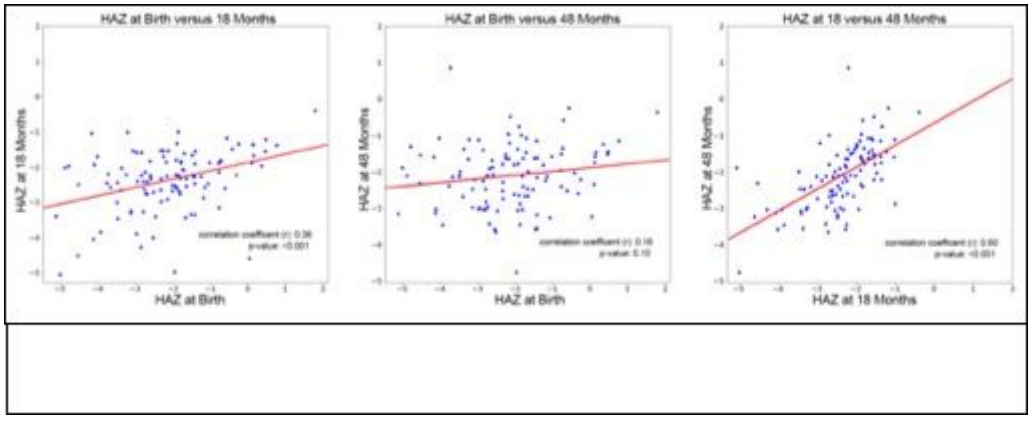

Figure 4

Comparison of Height for Age z-scores at birth, 18 months, and 48 months. Key: HAZ - Height for Age z-score

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- AdditionalFile1.xlsx