Methodological Developments in Growth Modeling

Stef van Buuren (stef.vanbuuren@tno.nl)

1. Netherlands Organization for Applied Scientific Research TNO, Leiden
2. Department of Methodology and Statistics, Utrecht University

Abstract

Appropriate modeling of human growth is needed to detect and counter poverty, poor health and nutrition, and lack of early stimulation. This presentation highlights three recent societal trends that impact on the statistical methodology of growth models. First, the spread of smartphones removes important barriers on information access for low and middle income countries (LMIC). A second trend is personalization of interventions for individuals. Finally, the impact of peers on growth is likely to be an important factor but has hardly been studied. This paper proposes curve matching as a technology that will help to signal abnormal growth and development in due time, to evaluate the effectiveness of interventions for a given child, and thus contribute to reduce the number of children who fail to reach their potential in cognitive development.

Key words: LMIC, curve matching, smartphone, big data, infant growth

1 Problem

More than 200 million children per year under five years fail to reach their potential in cognitive development because of poverty, poor health and nutrition, and lack of early stimulation.(Grantham-McGregor, 2007) These problems are related to deprivation, but it is not clear how to counter the impeding factors in a successful way.

The adverse effects of undernutrition, poverty and disease are particularly severe for infants and children under the age of two years, resulting in irreversible damage that extend far into their future lives. Stunting is the failure to achieve one's own potential for height. It is a symptom of the severe, irreversible physical and cognitive harm, initiated by chronic malnutrition, low hygiene, poverty and disease early in a child's life, often already present before birth. Reduction of stunting at the age of two years, both physical and cognitive, is an indicator of improved living conditions.

The Healthy, Birth, Growth and Development knowledge integration (HBGDki) program of the Bill & Melinda Gates Foundation targets the first 1000 days of life. The goal of the initiative is to determine what packages of interventions should be delivered to which group of individuals and at what age to reduce the burden of fetal growth impairment, preterm birth, stunted postnatal growth, and impaired cognitive development. The HBGDki aims to find out what works for whom, and when interventions be applied to result in an optimal growth and development during the first 1000 days.

The definition and measurement of physical and cognitive stunting, as well as the determination of timely and appropriate interventions, requires appropriate insight into the statistical regularities of childhood growth. In practice, a child's potential height needs to be defined, the prediction of future height should be made clear, the expected effect of an intervention should be estimated, and so on.

In this paper I will discuss three societal trends that impact upon the statistical technology needed to childhood growth: the rise of information technology, the increased interest in personalization, and the growing role of peer networks. I will then make a proposal that addresses the need emerging from these trends.

2 Trends

2.1 Rise of information technology

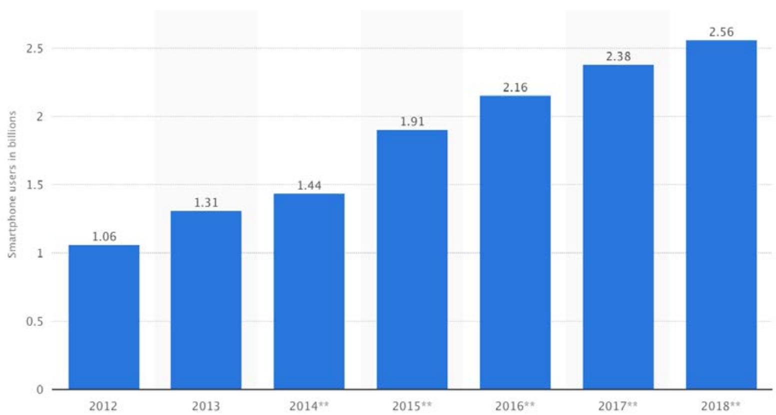

Figure 1: Number of smartphone users worldwide 2012 to 2018 (in billions) (source: Statistica 2015)

The first iPhone was introduced in the year 2007. The rate of adoption has been phenomenal. Figure 1 shows that nowadays about 2 billion people use a smartphone. The smartphone is now part of daily life. The number of users is expected to grow further to one in every three persons living on the planet in 2018.

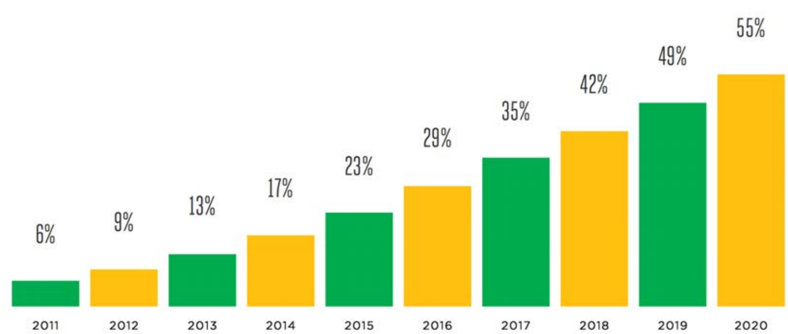

Figure 2: percentage of smartphone users relative to the total connections in Sub-Saharan Africa in the period 2011–2020. (Source: GSMA Intelligence)

Figure 2 illustrates that growth of smartphones is not restricted to the Western world. The smartphone is also omnipresent in poor areas of the world. This means that access to information in Low and Middle Income Countries (LMIC) needs no longer be a bottleneck.

2.2 Increased interest in personalization

In 2007, then senator Barack Obama proclaimed: *We are in a new era of the life sciences, but in no area of research is the promise greater than in personalized medicine*. In January 2015, this was followed by the launch of the Precision Medicine Initiative: *The time is right to unleash a new wave of advances in precision medicine*. Precision medicine is an emerging approach for disease prevention and treatment that takes into account people's individual variations in genes, environment, and lifestyle. The initiative is expected to generate the scientific evidence needed to move the concept of precision medicine into clinical practice.

Precision medicine connects very well to the daily practice of the health professional or physician. In practice, the health professional would like to be able to have answers on the following questions:

- Given what I know of the child, how will (s)he develop in the future?
- How certain am I of the child's future growth?
- If I do not intervene, will development be normal?
- If I do intervene, will the child's growth be normal, or healthy?

The parents of the child are interested in knowing the answers to questions like:

- Is the growth of my child normal?
- What can be done counter my child's inhibited growth?
- What is the prognosis if we do nothing?
- How certain is this prognosis?

Personalized estimates and interventions require a system for reasoning with individual data relative to healthy population and unhealthy (patient) populations. How far is the individual on becoming a patient, what can be done to prevent this, and when?

2.3 Role of peer networks

Social networks provide infrastructure for the spread of smoking, drinking, aggression, school achievement, physical activity, happiness, and other aspects of human behavior. The interest in the effects of peer network is rapidly rising, as evidence from scientific books appearing in the popular press (Christakis, 2009) (Pentland, 2014).

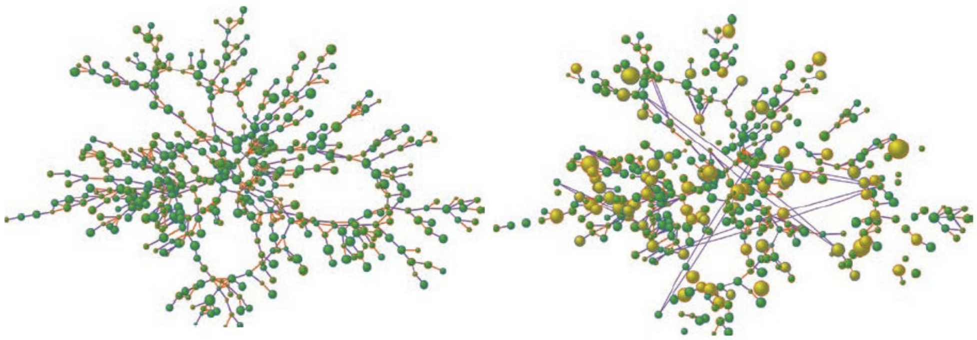

Figure 3: The social network of a Framingham heart study in 1975 (left) and 1990 (right). Each dot is a person. Social between two persons are indicated by lines. The color of the dots indicated whether the person is obese (green is non-obese, yellow is obese).

In a well-publicized study, Christakis and Fowler suggested that obesity should be treated as a contagious disease. (Christakis, 2007) Figure 3 illustrates the development of obesity within a social network during a period of 15 years. The salient feature is that the appearance of green dots occurs in clusters. This has three possible explanations: we choose one similar to use in our network, obesity is 'contagious' or, finally, there may be common confounding factors that structure both the network and obesity.

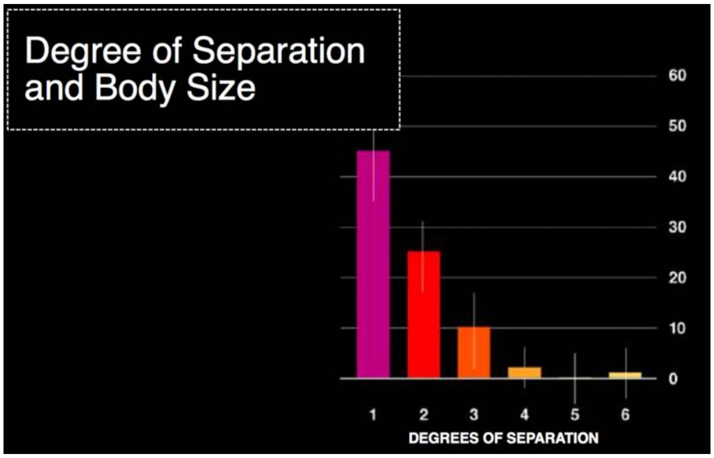

Figure 4: The probability that a person is obese given that a social contact is obese (%).

Christakis and Fowler have gone to great lengths in arguing for the contagious explanation. Figure 4 was presented by them as a way to show how much your friend's, your friend's friend's obesity affect your probability of being obese. It is only in the 5th degree that there is no influence anymore. Others have disputed the exclusive focus on the contagion hypothesis, the most severe criticisms being voiced by Lyons.(Lyons, 2011) Clustering of obesity itself is undisputed, but it is yet unclear as to what mechanism leads up to the phenomenon.

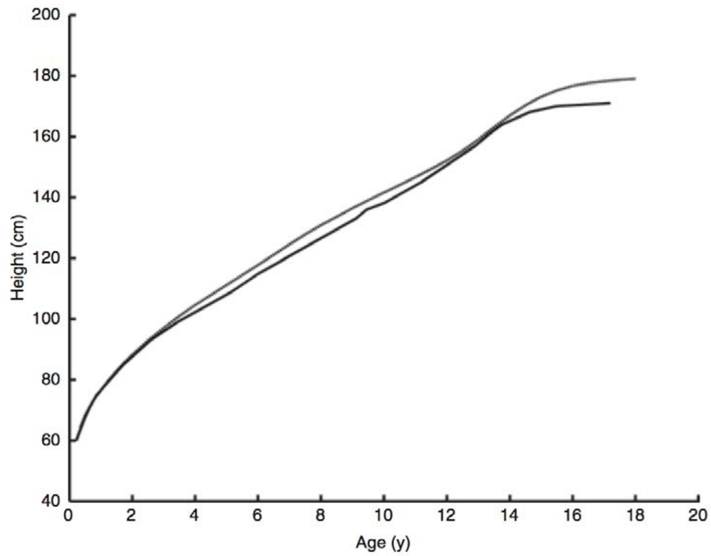

Figure 5: Body height of German boys measured in 1893 (dark line) and 2011 (grey line). Age is given as biological age with tempo adjustment via peak height velocity mapping. Source: (AßmannHermanussen, 2013)

Are there also peer effects on final height? Figure 5 shows that – when corrected for differences in tempo – height differences between two German born apart 120 years are initially small, but markedly increase during adolescence. Since the period of adolescence coincides with the forming of social networks, Aßmann and Hermanussen suggest the existence of a "community-based target" in height that might explain the increase during adolescence. (Aßmann, 2013) In other words, you will become if your friends are taller. The idea is intriguing, but it is yet an open question whether it is actually true.

2.4 Impact

This section highlighted three societal trends that impact on the methodology needed to model childhood growth: the rise of information technology, the increased interest in personalization and the role of peer networks. Accommodating for these trends implies that our statistical models need to answer question on individual persons, located within the context of their social network, about the effects of potential interventions applied either directly to the individual, or to the social network around the individual. The results should be made available immediately and everywhere, and in a form that everyone can understand.

3 A proposal

3.1 The idea

The wide availability of smartphones in LMIC opens up new ways of providing, monitoring and enhancing preventive child health care during the first 1000 days. This section introduces an internet-based system that aims to educate and help those who need it most.

Let us consider first the ideal scenario. Suppose a family visits their family doctor with their newborn child, with a concern whether the child is healthy. The family doctor asks for specific complaints, measures and interprets height, weight and head circumference of the child, and inspects the child for signs of dehydration, jaundice, malnutrition, malaria or other common health problems in infants. The doctor may then inform the parents, and take appropriate action.

The problem with this scenario is that in many LMIC there simply is no family doctor. The doctor might live too far away, could be too expensive, might be culturally unacceptable, or may have inadequate resources for help. The consequence: symptoms are not being diagnosed in time, and interventions will be too late, resulting in now irreversible physical and cognitive damage to the child, with possibly life long consequences.

The ‘next best’ scenario is to replace the function of the family doctor by a less educated care taker, assisted by an intelligent system that knows all about child growth and its associated perils. Such care takers can educate themselves when they have access to the internet by a smartphone, tablet or desktop. The care taker discusses the complaints with the care taker, possibly guided by the remote system, measure and input relevant parameters into the system, and interpret the output of the system with the parents, preferably accompanied by appropriate therapeutic actions. The ambition is that this develops into a low-cost high-quality standardized health service for many.

3.2 An example

Figure 6: Snapshot of the Dutch growth prediction system based on curve matching.

Growth monitoring aims to detect pathology in time, so that adequate treatment is (still) possible. Figure 6 displays a web application currently being developed for The

Netherlands. The shaded area in Figure 6 portrays the normal variation in head circumference. The system aims to predict growth, and detect abnormal growth before it actually happens.

Technically, the system is based on curve matching. The idea is as follows. Suppose that Laura is three months old, that we have measured her head circumference at months 1, 2, and 3, and that we have plotted the data as the red curve in Figure 6. Suppose we aim to predict, for baby Laura, her head circumference at the age of 14 months (the outcome). Figure 6 plots the data of Laura, together with the data of 10 matched children from our database of existing children ”who are like Laura”. There are many ways to express similarity, but here it is defined in terms of the difference between the predicted values from the regression of head circumference up to month 3 (predictors) on head circumference at 14 months. The grey curves suggest how Laura’s future growth curve may look. Since the grey curves are close together, we can be reasonably confident about Laura’s future head growth. The blue curve summarizes the 10 grey curves in the most likely future trajectory.

Now suppose that we wish to determine the effect of providing malaria bed nets on the head circumference of Laura at 14 months. We will repeat the analysis, but with one important difference. We first split the donor data into two groups: Group A is a group that did obtain malaria bed net before the age of 3 months, whereas Group B did not receive bed nets. Matching is now done in both Group A and Group B separately, resulting in two sets of grey curves. If both bundles of matched curves are similar, then obtaining bed nets have little or no effect on future head circumference. However, when the bundles diverge, e.g. if those for Group are consistently lower, than handing out a bed net to Laura is an effective way to increase her future head circumference, potentially leading to improved cognition and chances in life. Note that the estimate applied only to Laura. The result may be different for other children, especially if they differ from Laura.

Note that the method just described provides a way to organize – when considered ethical – embedded trials. We may first calculate the potential benefit of the intervention, and then randomize the intervention only to those children within the subgroup of potential benefit. In this way, an efficient use of resource is possible.

More details on curve matching can be found elsewhere.(van Buuren, 2014)

4 Conclusion

This paper discusses three societal trends that impact upon the statistical technology needed to childhood growth: the rise of information technology, the increased interest in personalization, and the growing role of peer networks. The development of the statistical models in this paper was framed to improve living condition in LMIC. The method should be designed to answer question on individual persons, located within the context of their social network, and about the effects of potential interventions applied either directly to the individual, or to the social network around the individual. The results of the modeling should be made available immediately and everywhere, and in a form that everyone can understand.

I believe that the curve matching technology as illustrated in section 3 can be helpful to signal abnormal growth and development in due time, to evaluate the effectiveness of interventions for a given child, and thus contribute to reduce the number of children who fail to reach their potential in cognitive development.

I realize there are still a lot of open methodological issues. We need asymptotic theory for curve matching, we need to develop theory for estimating individual causal effects by matching, and we need to develop techniques for incorporating peer effects in multivariate data.

Notwithstanding these difficulties, developments like these offer not only a systematic way to educate care takers in LMIC, but also provide a handle on efficient experiments into the effects of interventions aimed at improving chances to live and healthy and productive life.

Acknowledgement

This study was kindly supported by the Bill & Melinda Gates Foundation.

References

Aßmann, C., & Hermanussen, M. (2013). Modeling determinants of growth: evidence for a community-based target in height. *Pediatric research*, *74*(1), 88-95. Retrieved from http://www.nature.com/pr/journal/v74/n1/abs/pr201350a.html

Christakis, N. A., & Fowler, J. H. (2009). *Connected. The Surprising Power of Our Social Networks and How They Shape Our Lives - How Your Friends' Friends' Friend Affect Everything You Feel, Think, and Do*. New York: Back Bay Books.

Christakis, N. A., & Fowler, J. H. (2007). The spread of obesity in a large social network over 32 years. *New England journal of medicine*, *357*(4), 370-379. Retrieved from http://dash.harvard.edu/bitstream/handle/1/3710802/Christakis_SpreadofObesity.pdf?sequence=2

Grantham-McGregor, S., Cheung, Y. B., Cueto, S., Glewwe, P., Richter, L., Strupp, B., & Group, I. C. D. S. (2007). Developmental potential in the first 5 years for children in developing countries. *The lancet*, *369*(9555), 60-70. Retrieved from http://www.sciencedirect.com/science/article/pii/S0140673607600324

Lyons, R. (2011). The spread of evidence-poor medicine via flawed social-network analysis. *Statistics, Politics, and Policy*, *2*(1). Retrieved from http://arxiv.org/pdf/1007.2876

Pentland, A. (2014). *Social Physics: How Good Ideas Spread-The Lessons from a New Science*. Penguin. Retrieved from http://books.google.com/books?hl=en&lr=&id=vcgqAAAAQBAJ&oi=fnd&pg=PT7&dq=pentland+social+physics&ots=4pFlq9NYt4&sig=oFfCG5Ef5OCF22jO4OZhb22V04I

van Buuren, S. (2014). Curve matching: A data-driven technique to improve individual prediction of childhood growth. *Annals of Nutrition & Metabolism*, *65*(3), 227-233. doi:DOI: 10.1159/000365398