MAX PLANCK INSTITUTE
FOR EVOLUTIONARY ANTHROPOLOGY

STUDY INFORMATION AND DATA USE AGREEMENT FOR THE RESEARCH PROJECT: "*CROSS CULTURAL COMPARISON OF HUMAN GROWTH TRAJECTORIES*"

Catalina I. Fernández, Caissa Revilla-Minaya & John A. Bunce

Max Planck Institute for Evolutionary Anthropology,
Deutscher Platz 6, 04103 Leipzig, Germany

1. Brief Description

In this project, we will apply the growth model recently developed by Bunce, Fernández, and Revilla-Minaya 2022 (see: Causal models of human growth and their estimation using temporally-sparse data) to longitudinal datasets from populations inhabiting a wide range of ecoregions with the aim of distinguishing and describing general patterns and global variation in human growth trajectories. By using this model, we expect to be able to estimate, and hopefully disentangle, the relative contributions of different factors affecting growth in height and weight around the world, as well as to compare specific components of growth among diverse populations.

2. Intended use of the Data

The data provided by the collaborators in this project will be used for the sole purpose of this study on the comparison of growth trajectories. The final product will be one or more scientific articles on this topic. Data will not be used for any commercial purposes. We will not distribute these data to third parties under any circumstance. We will make no attempt to identify or contact research participants, their households or communities. Collaborators will have the option of making the de-identified data they contribute publicly available as an accompaniment to the published article(s) that result from this project (see below).

3. Dataset and population (s)

Please provide the following characteristics of the dataset (s), indicating population, location (s), a brief description of data collection methods and/or a reference for them, and indicate whether data was collected after IRB approval or other ethics review (please, indicate review board and protocol number). If other ethical protocols were followed, please specify.

- **Population name(s) and brief description (e.g., subsistence practices, refugees from another region, access to Western healthcare systems, common health challenges, etc.):**

The Indigenous Xavante are among the larger ethnic groups in Brazil, with a population of about fifteen thousand living in nine Indigenous reserves. Like many other Gê-speaking Indigenous groups in Brazil, the Xavante were considered semi-nomadic by early observers because they depended more on gathering and hunting than gardening during most of the year. After prolonged contact with non-Indigenous Brazilian society and confinement within relatively small Indigenous reserves, particularly since the late 1960s, the Xavante subsequently became increasingly sedentary and came to depend more on agriculture for food production. The populations now residing in Pimentel Barbosa and Etênhiritipá villages settled in their current approximate locations in the early 1970s due to pressures by governmental agencies to vacate their previously expansive territory in order to facilitate settlement by large-scale. Currently, they also depend on market resources to meet dietary and other needs. Sugar, coffee, salt, pasta, soft drinks, crackers, and cooking oil are now regular items in the typical Xavante diet. In addition, they now have access to increasingly diverse sources of monetary income, including salaried employment and social benefits, which further contribute to reliance on purchased food items.
You can see more details in:

Coimbra et al. 2002 - https://www.researchgate.net/publication/346945618_The_Xavante_in_Transition_Health_Ecology_and_Bioanthropology_in_Central_Brazil
Welch et al. 2009 - https://link.springer.com/article/10.1007/s10745-009-9216-7
Welch et al. 2013 - https://acervo.socioambiental.org/acervo/livros/na-primeira-margem-do-rio-territorio-e-ecologia-do-povo-xavante-de-wedeze

- **Location (s) (please be as specific as possible, and include GPS coordinates if available):**

The Xavante indigenous land, where all the Xavante live, located in the east of the State of Mato Grosso, in Central Brazil. No GPS coordinates available. But in Welch et al. 2013 (https://acervo.socioambiental.org/acervo/livros/na-primeira-margem-do-rio-territorio-e-ecologia-do-povo-xavante-de-wedeze), have more details. Pimentel Barbosa Indigenous Reserve, local of study, is the largest Xavante reserve, with an area of 328,966 ha. The study villages, Pimentel Barbosa and Etênhiritipá, at the time of the research constituted more than half of the Xavante indigenous people in the reserve.
The tropical landscape is dominated by cerrado vegetation and has average annual temperatures of 20-22ºC. Although the cerrado, a tropical savanna landscape, shares some biological and

ecological features with Amazonian tropical rainforests, is has considerably greater seasonal variation in temperatures and rainfall, greater overall biodiversity, and more geographically dispersed plant and animal resources.

- **Data collection methods (please include reference if available):**

The results presented here are based on six data collection waves conducted in July 2009, January and July 2010, January and July 2011, and January 2012. All children under 10 years of age resident in the two villages during any collection wave were allowed to enter or exit the study at any time. Possible reasons for a child entering the study after the first wave or exiting before the last included birth, death, relocation to or from the study villages, and parents' decision to not participate in one or more waves. Additionally, a child was automatically unenrolled from the study upon reaching 10 years of age. The number of measurements varied through time (unbalanced data), as did the intervals between data points for some children (temporally unstructured data). Thus, different numbers of children were evaluated the maximum of 6 times (during all data collection waves) through the minimum of one time (during only one wave).
Birthdates were obtained from healthcare records in possession of parents or guardians or, alternatively, local health services (a village health post). Birth weight data were also obtained from local health service records, but were excluded from the diachronic dataset used to construct growth curves because they were unavailable for children born before 2008 and did not follow our measurement protocol.
Anthropometric data (weight and height) were sought for all resident children of both sexes < 10 years of age and physically capable of being measured and weighed. Participants' weight and height measurements were collected during home visits by three previously trained and standardized researchers following Lohman et al. (1988). Children ≤ 24 months were measured in the recumbent position to 0.1 cm accuracy using Seca 416 (Hamburg, Germany) pediatric anthropometers. Older children were measured in the standing position to 0.1 cm accuracy using Seca 214 (Hamburg, Germany) anthropometers. Weight was measured to 0.1 kg accuracy using Seca 872 (Hamburg, Germany) portable electronic scales, which include a mother/child function. For all measures, children were asked to be barefoot and wear minimal clothing.
This information was taken in full from the article Ferreira et al. 2016 - https://pubmed.ncbi.nlm.nih.gov/27239686/
Other information can be found at: https://pubmed.ncbi.nlm.nih.gov/31654538/

- **IRB/ethical review board approval (yes/no, explain):**

The present study occurred within larger multidisciplinary research projects addressing diachronic health transition among the Xavante of Pimentel Barbosa Indigenous Reserve. The study was approved by the Research Ethics Committee at the National School of Public Health (Escola Nacional de Saúde Pública), Oswaldo Cruz Foundation (Fundação Oswaldo Cruz), and the National Research Ethics Council (Conselho Nacional de Ética em Pesquisa). Permission to conduct the study within the Indigenous reserve was authorized by the National Indian Foundation (Fundação Nacional do Índio). The project was presented to residents and leaders of Pimentel Barbosa and Etênhiritipá villages during community meetings. Village leaders signed a

Collective Prior Informed Consent on behalf of residents. Parents and guardians were allowed to decline participation on behalf of their children and were present for anthropometric measurement.

4. Full name, affiliation and e-mail of all contributors/collaborators for this dataset (s):

Aline Alves Ferreira – alineaf@nutricao.ufrj.br
Instituto de Nutrição Josué de Castro, Universidade Federal do Rio de Janeiro, Avenida Carlos Chagas Filho 373, Bloco J. Rio de Janeiro, RJ 21941-902, Brazil.

James R. Welch
Carlos E. A. Coimbra Jr

Escola Nacional de Saúde Pública, Fundação Oswaldo Cruz, Leopoldo Bulhões 1480. Rio de Janeiro, RJ 21041-210, Brazil.

.

5. Data submission

Please send a **link to a folder** containing a file or several files (in case of more than one dataset) in **CSV format** to catalina_fernandez@eva.mpg.de.

- Please include the **name of the population in the file name**, and an additional dataset identifier in case of multiple datasets.
- If the dataset also contains single (as opposed to longitudinal) measures for some individuals, please do not remove them from the file. We will also use those observations in the analysis.
- If you have data on adult individuals, whether or not they were also measured as children/adolescents, please include these observations as part of your dataset.

Please, do not include any direct identifying information in the dataset relating to the individuals, their relatives or household members, such as names, addresses, telephone numbers, e-mail addresses or social media identifiers, etc.

For each dataset, we request collaborators on this project to share a **deidentified** data file (.csv) containing the following information and columns:

Column 1: Individual identification code or number that is consistent across longitudinal observations of the same individual.
Column 2: Date of birth in **dd/mm/yyyy** format.
Column 3: Date of data collection event in **dd/mm/yyyy** format.
* If date of birth is not available, please provide **age** in days, whenever possible, for each data collection event.
* If known age is uncertain, if possible, please provide a **range** of minimum, estimated, and maximum age, based on your best approximations.
Column 4: Sex; coded as **m** or **f**.
Column 4: Height in cm
Column 5: Weight in kg

*Note that, for longitudinal measures, you will have multiple rows with the same individual identification code, date of birth, and sex, but with different dates of data collection, heights, and weights.

6. Long term archiving and data access

One of the objectives of this project is to make the deidentified datasets and code (s) used in the data cleaning and analyses available in an open-access repository. We encourage all collaborators to commit to share and deposit deidentified data in a public repository created for this project, once they have checked that, by doing so, they are not violating compliance with their ethical review protocols (e.g., IRB) or other agreements with research participants and community members/leaders. Making data open-access is not a requirement to participate in this study, but we believe that by granting public access to the code and data used we will allow other researchers to check our analyses and reproduce our findings.
If you agree that data from your research site can be made publicly available for this purpose, we will create an appropriate repository (e.g., on Github) to curate this dataset and indicate the contact information for the researchers responsible for each field site, in case someone wishes to use the dataset for a purpose other than simply checking the results of our analysis. If for any reason, in the future, you wish to change your decision regarding data access and usage, we can either include or remove the data from this repository.

Use of the data and access on a public repository (choose one option):

_____ **Restricted use**. I agree that deidentified data can be made publicly available for the sole purpose of reproducing this specific analysis. Data cannot be used by the scientific community or the public for any other purpose.

_____ **Unrestricted**. I agree that deidentified data can be made publicly available without barriers to access or use.

____**x**__ **I do not consent to making this data available in a public repository**. I am sharing this data only for the purpose of this study. I do not agree that this dataset(s) or any portion of it/them can be shared publicly under any circumstances.

7. Data sharing information to the communities involved

We believe that one of our responsibilities as scientists is to communicate research findings with the communities and individuals who provided their time and biometric information in order to make this study possible. We are aware that most, if not all, collaborators on this project have active field sites and engage regularly with the communities they work with for the purpose of informing and sharing research findings and other initiatives alike. We ask collaborators to this project to share the results of this study with the contributing communities, particularly the parts that concern the specific population that they work with. If a collaborator is no longer in contact with the study population, please let the project leaders know so that together we can potentially brainstorm an alternative solution. Ideally, presenting results to participants will occur prior to publication, so that participating communities have the opportunity to (re)express their permission for us to publish the results. Sharing the results may take the form of a live Power Point presentation, a video, a written information sheet, or any other format that the collaborator believes is most culturally appropriate to present this kind of information to the communities and participants who provided the data. If it is of interest as a model or guide, the project leaders (Catalina Fernández, Caissa Revilla-Minaya and John Bunce) can share with the contributors the materials that they will design for the Matsigenka population that they work with.

8. Manuscript authorship

The project leaders (Catalina Fernández, Caissa Revilla-Minaya and John Bunce) will draft the main manuscript and decide the order of authorship and co-authorship. We will ask all collaborators for feedback once analyses are completed and while we draft the discussion. We are currently considering organizing a workshop at MPI in Leipzig for project collaborators, after analyses are completed, but prior to publication, in order to share the results of the analyses, discuss the main findings and potential causal mechanisms contributing to variation in growth in different populations. This can also be an opportunity for planning further analyses and related projects for the future. More information about this potential meeting/workshop will be shared in the next months.

9. Tentative timeline

- 30 June 2023: All datasets have been submitted to the project leaders.
- 30 September 2023: All datasets have been checked for errors and formatted for analysis.
- 30 March 2024: Data analysis is complete; manuscript is drafted by project leaders.
- Mid-May 2024: Potential workshop for contributors/collaborators at MPI Leipzig.
- 30 July 2024: Results have been presented to participating communities.
- 30 August 2024: Main manuscript is submitted for publication.

***Note that this is a very tentative timeline.**

10. Questions

For questions regarding data submission, data archiving, and other aspects of this study, or additional information, please contact Catalina Fernández by e-mail at catalina_fernandez@eva.mpg.de.

Name of the person in charge of the dataset: Aline Alves Ferreira (alineaf@nutricao.ufrj.br)

Date: July 09, 2023