Plan Overview

A Data Management Plan created using DMPTool

DMP ID: https://doi.org/10.48321/D1ZM0V

Title: DMSP for "An interdisciplinary approach to improve age estimates in biological anthropology"

Creator: Sara Casado zapico - ORCID: 0000-0002-3605-5690

Affiliation: New Jersey Institute of Technology

Principal Investigator: Leticia Rubio Lamia, Sara Casado Zapico

Data Manager: Francisco Medina-Paz, Cristo León

Contributor: Francisco Medina-Paz, Stella Martin-de-las-Heras

Funder: National Science Foundation (nsf.gov)

Funding opportunity number: NSF23-503


Template: NSF-SBE: Social, Behavioral, Economic Sciences

Project abstract:

Age-at-death estimation is one of the key components to establish the biological profile of an individual in bioanthropological contexts. In subadults this is quite precise because growth and development normally occur at a predictable rate. In contrast in adults is less precise because growth has ceased, age estimation is based on degenerative changes of bone and teeth and it can be affected by endogenous and exogenous factors, lifestyle, etc.. Although there have been many attempts to improve age estimation in adult individuals based on degeneration of body parts, at best, these methods accomplish estimation of physiological age, which may significantly differ from chronological age. As a result, it is not uncommon for age-at-death estimates of adult individuals to be presented as an age range spanning 10 years or more. New methodologies for age estimation are required. The goal of this proposal is to improve age-at-death estimation in bioanthropological contexts through applying an interdisciplinary approach, based on the biochemical and physicochemical bases of aging and studying the hardest tissues in human body, teeth. To address this, MicroCT, XRD and FTIR will be
performed to assess the internal physicochemical changes in tooth structure and their relationship with the age. At the biochemical level, mtDNA mutations, methylation markers and proteomic analyses will be evaluated to correlate their changes with aging. Statistical analyses will be carried out to identify the physicochemical and biochemical changes highly associated with the age, leading to the development of different prediction models for age estimation based on these techniques alone or in combination. The experimental and interdisciplinary approach of this proposal will provide a full array of physicochemical and biochemical techniques with their accuracies for age estimation (+/- 1-3 years actual age), and their feasibility to apply in different bioanthropological contexts, towards a significant improvement to the overall precision of individual biological profiles.

Start date: 01-01-2024

End date: 12-31-2026

Last modified: 07-25-2023

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customize it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal.
DMSP for "An interdisciplinary approach to improve age estimates in biological anthropology"

Roles and responsibilities

The DMP should outline the rights and obligations of all parties as to their roles and responsibilities in the management and retention of research data. It should also consider changes to roles and responsibilities that will occur should a principal investigator or co-PI leave the institution or project. Any costs should be explained in the Budget Justification pages.

Dr. Sara C. Zapico (PI): Molecular biology (mtDNA and epigenetics) and statistical analyses. Coordination of the project, data gathering, article preparation and meeting presentations.

Dr. Leticia Rubio Lamia (CO-Is): Physicochemical analyses of the teeth applying micro-CT, XRD and ATR-FTIR. Hosting the graduate student, during her short-stay. Coordination of shipping of the teeth samples, data gathering, article preparation and meeting presentations.

Dr. Stella Martin de las Heras (CO-Is): micro-CT of the teeth and article preparation.

Dr. Hao Chen: Proteomics (CO-Is), coordination of data gathering, article preparation and meeting presentations.

Dr. Francisco Medina Paz (PM): Molecular biology part of the project, data management, and statistical analyses.

María Josefina Castagnola (DM): data analysis of microCT, XRD, and FTIR.

Cristo León (Other): Compliance.

Principal Investigator (PI): For the purposes of this document, the term PI refers to the proposal team, which may consist of an individual PI or a PI and Co-Is named in the proposal.

Co-Investigators (Co-Is): Co-Is have the same data-access rights as the PI, although the PI and Co-Is are expected to coordinate their efforts.

Project Manager (PM): For the purposes of this document, the term PM refers to the individual(s) coordinating the DMP administration.

Data Manager (DM): For the purposes of this document, the term DM refers to the individual(s) coordinating the cleaning and validation of data.

Data Management Plan (DMP): The DMP is the document that is a required part of the proposal to the U.S. National Science Foundation. It describes how a proposal will follow NSF policy on managing, disseminating, and sharing research results. General guidelines for data management plans are explained in PAPPG II.D.2(ii)[1].


Expected data

The DMP should describe the types of data, samples, physical collections, software, curriculum materials, and other materials to be produced in the course of the project. It should then describe the expected types of data to be retained.

For mtDNA and DNA methylation data: melting curves (qPCR) images: .jpeg, .tiff and .pdf format, and
pyrosequencing data FASTA files.

For micro-CT, FTIR and X-Ray diffraction data: .tiff and .csv and .xls* format files.

For proteomics data: .*ML format compressed files. Metadata output will be converted to .csv and .pdf format files. Homemade developed scripts will be converted to .txt format files.

**Period of data retention**

SBE is committed to timely and rapid data distribution. However, it recognizes that types of data can vary widely and that acceptable norms also vary by scientific discipline. It is strongly committed, however, to the underlying principle of timely access, and applicants should address how this will be met in their DMP statement.

All generated data will be stored for the whole duration of the project (3 years) from 2024-2026. After that time, published raw data, metadata, and homemade scripts will be submitted to the correspondent databases and repositories and virtually stored indefinitely. All data storage procedures will comply with European, national and NJIT regulations and security and privacy standards. Data only will be published in the form of articles or as supplementary material to the publications.

As per the OMB Circular A-110 (now 2 CFR, Ch. II, §215.36(d)(2)(i), and codified in 5 U.S.C. 552(a)(4)(A)) and federal public access (NSF SBE, 2016; Smith et al., 2021) our research data is NOT trade secrets nor Personnel and medical information.

**Sources**


**Data format and dissemination**

The DMP should describe data formats, media, and dissemination approaches that will be used to make data and metadata available to others. Policies for public access and sharing should be described, including provisions for appropriate protection of privacy, confidentiality, security, intellectual property, or other rights or requirements. Research centers and major partnerships with industry or other user communities must also address how data are to be shared and managed with partners, center members, and other major stakeholders.

For a broad dissemination of the research results and their relevance to the scientific community, the PI will present at meetings of the American Academy of Forensic Science (AAFS) and American Association of Physical Anthropology (AAPA), and she will continue seeking out opportunities to conduct technical workshops for professionals in the field. Because the product of the proposed research is relevant to a broad scientific audience,
submission of research findings will not be limited exclusively to specialized peer-reviewed journals. The results of this research, journals of a more generalized nature (molecular biology, analytical chemistry, genomics, medicine) will be considered as well. To facilitate widespread use of this methodology, the PI plans to partner with Medical Examiner’s Offices around the United States. Additionally, The Argentine Forensic Anthropology Team (EAAF) has expressed interest in learning and applying biochemical-based age-at-death estimation tools to their worldwide cases. Furthermore, in order to make an impact on future generations of scientists, this project will be integrated with undergraduate and graduate students training, mentoring and curricula.

Data storage and preservation of access

The DMP should describe physical and cyber resources and facilities that will be used for the effective preservation and storage of research data. These can include third party facilities and repositories.

All data (except for the one published in research articles and congresses) produced at University of Malaga will be kept for 5 years, after which it is destroyed. Data will be stored in Google Drive, which is a system that has level 3 security and protection at the University of Malaga and also on physical hard drives that can only be accessed with a password (owned by the project researchers) and are kept under lock and key. Google Drive detects and blocks threats, phishing, and malware. When published, raw data and metadata, as well as homemade scripts will be submitted to NCBI database and Github repository, respectively.

Additional possible data management requirements

More stringent data management requirements may be specified in particular NSF solicitations or result from local policies and best practices at the PI’s home institution. Additional requirements will be specified in the program solicitation and award conditions. Principal Investigators to be supported by such programs must discuss how they will meet these additional requirements in their Data Management Plans.

A postdoctoral mentoring plan will be submitted with the proposal.

An international data-sharing agreement with "The University of Málaga" will facilitate collaboration.

As a member of both the DNA and Pathology sections of the Interpol Disaster Victim Identification Group, the PI intends to share the results of the proposed research with the consortium and foster discussion about the potential standardization of these methodologies in the field. To facilitate the widespread use of this methodology, the PI plans to partner with Medical Examiner’s Offices around the United States. Additionally, The Argentine Forensic Anthropology Team (EAAF) has expressed interest in learning and applying biochemical-based age-at-death estimation tools to their worldwide cases, thereby promoting the exchange of knowledge as well as increasing the project’s overall impact and significance.

For a broad dissemination of the research results and their relevance to the scientific community, the PI will present at meetings of the American Academy of Forensic Science (AAFS) and American Association of Physical Anthropology (AAPA), and she will continue seeking out opportunities to conduct technical workshops for professionals in the field.
Because the product of the proposed research is relevant to a broad scientific audience, submission of research findings will not be limited exclusively to specialized peer-reviewed journals. Although peer-reviewed journals related to anthropology or forensics are certainly appropriate venues for sharing the results of this research, journals of a more generalized nature (molecular biology, analytical chemistry, genomics, medicine) will be considered as well. The nature of the proposed work will inherently contribute to medical research, particularly with respect to gerontology. By analyzing different biochemical and physicochemical parameters on dental tissues to identify age-related changes, the proposed project creates a framework of reference data that advances research into novel targets of aging.

To facilitate widespread use of the validated tool for accurate age-at-death estimation, the PI plans to partner with Medical Examiner’s Offices around the United States. Additionally, The Argentine Forensic Anthropology Team (EAAF) has expressed interest in learning and applying biochemical-based age-at-death estimation tools to their worldwide cases, thereby promoting the exchange of knowledge as well as increasing the project’s overall impact and significance.
Planned Research Outputs

Text - "TDB peer-review articles"

Preparation individual article Micro-CT - 2025
Preparation individual article XRD and FTIR - 2025
Preparation individual article mtDNA - 2026
Preparation individual article proteomics - 2026
Preparation individual article epigenetics - 2026
Preparation of a global article combined physicochemical and biochemical methodologies and findings - 2026

Event - "Presenting finding at congresses"

Presenting the micro-CT, XRD and FTIR findings at the congress of the Spanish Association of Forensic Anthropology and Odontontology - 2025
Presenting the findings at the American Academy of Forensic Science meeting 2026
Presenting the micro-CT, XRD and FTIR findings at the International Academy of Legal Medicine - 2026
Presenting the findings at the American Association of Biological Anthropology meeting 2026

Interactive resource - "Micro-computed tomography"

High resolution 3D X-ray images

Workflow - "Age estimation model"

A detailed sequence of steps in order to use the validated age-estimation model properly according to the biochemical and/or physicochemical data available.

This information will be shared with the DNA and Pathology sections of the Interpol Disaster Victim Identification Group and foster discussion about potential standardization of these methodologies in the field.

Dataset - "Attenuated total reflectance Fourier Transform Infrared Spectroscopy"

Analysis and quantification of Crystallinity index, Mineral-organic matrix ratio, amides I, Carbonate substitution and Collagen maturity.

Dataset - "Shotgun Liquid Chromatography tandem Mass Spectrometry profiles"

Protein dataset from LC/MS-MS analysis focused on collagens, pentosidines, non-collagenous proteins, amelogenic proteins, fibrillins, adhesion proteins, and glycoproteins.

Dataset - "Protein quantitation"

Label-free quantitation of structural and functional proteins based on ion intensities, and absolute quantitation of
selected proteins with Coulometric Mass Spectrometry.

**Dataset - "DNA methylation patterns"**

Pyrosequencing data from KLF14, NPTX2, ELOVL2, FHL2, SCGN, PDE4C, EDARADD, and other genes currently under research.

---

**Planned research output details**
<table>
<thead>
<tr>
<th>Title</th>
<th>Type</th>
<th>Anticipated release date</th>
<th>Initial access level</th>
<th>Intended repository(ies)</th>
<th>Anticipated file size</th>
<th>License</th>
<th>Metadata standard(s)</th>
<th>May contain sensitive data?</th>
<th>May contain PII?</th>
</tr>
</thead>
<tbody>
<tr>
<td>TDB peer-review articles</td>
<td>Text</td>
<td>2026-12-11</td>
<td>Open</td>
<td>Open Science Framework Digital Commons NJIT</td>
<td></td>
<td>Creative Commons Attribution Non Commercial No Derivatives 4.0 International</td>
<td>Dublin Core PREMIS</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Presenting finding at congresses</td>
<td>Event</td>
<td>2026-12-11</td>
<td>Open</td>
<td>None specified</td>
<td>None specified</td>
<td>None specified</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Micro-computed tomography</td>
<td>Interactive resource</td>
<td>2024-12-30</td>
<td>Open</td>
<td>None specified</td>
<td>None specified</td>
<td>None specified</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Age estimation model</td>
<td>Workflow</td>
<td>2026-12-19</td>
<td>Open</td>
<td>None specified</td>
<td>None specified</td>
<td>None specified</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Attenuated total reflectance Fourier Transform Inf ...</td>
<td>Dataset</td>
<td>2025-06-29</td>
<td>Open</td>
<td>None specified</td>
<td>None specified</td>
<td>None specified</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Shotgun Liquid Cromatography tandem Mass Spectrome ...</td>
<td>Dataset</td>
<td>2025-12-30</td>
<td>Open</td>
<td>None specified</td>
<td>None specified</td>
<td>None specified</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Protein quantitation</td>
<td>Dataset</td>
<td>2025-12-19</td>
<td>Open</td>
<td>None specified</td>
<td>None specified</td>
<td>None specified</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>DNA methylation patterns</td>
<td>Dataset</td>
<td>2026-06-19</td>
<td>Open</td>
<td>None specified</td>
<td>None specified</td>
<td>None specified</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>