Data Management and Sharing Plan

Element 1: Data Type:

A. Types and amount of scientific data expected to be generated in the project:

Summarize the types and estimated amount of scientific data expected to be generated in the project.

The proposed study is expected to generate a significant amount of clinical, laboratory, and imaging data from approximately 360 participants (180 mother-infant pairs) across three sites. The data will include but not be limited to: medical history, demographic information, patient characteristics, lab tests (urinalysis, analysis of angiogenic factors, uric acid, glycemia), physical exams, blood pressure measurements, fetal monitoring data (estimated fetal weight, non-stress tests), pregnancy duration, neonatal outcomes (birth weight, Apgar scores, head and chest circumference at birth, NICU admission), anthropometric measurements, neurodevelopmental assessments (Bayley scale of infant development at 3, 6, and 12 months), and various imaging data (2D Doppler studies of uterine, umbilical, and middle cerebral arteries, 3D images of the fetal thigh, real-time grayscale 2D clips of the fetal heart).

B. Scientific data that will be preserved and shared, and the rationale for doing so:

Describe which scientific data from the project will be preserved and shared and provide the rationale for this decision.

All collected data will be preserved and shared to maximize transparency, reproducibility, and the potential for future research collaborations and discoveries. This includes all clinical, laboratory, and imaging data listed above. The rationale is to contribute meaningfully to the scientific community's understanding of early onset preeclampsia and its management, potentially leading to improved maternal and infant outcomes.

C. Metadata, other relevant data, and associated documentation:

Briefly list the metadata, other relevant data, and any associated documentation (e.g., study protocols and data collection instruments) that will be made accessible to facilitate interpretation of the scientific data.

Associated metadata will include but not be limited to study protocols, data collection instruments, consent forms, and detailed descriptions of the methodologies used for data collection and analysis. Other relevant data may encompass quality control measures, data validation processes, and any software or algorithms used in the analysis of imaging data.

Element 2: Related Tools, Software and/or Code:

State whether specialized tools, software, and/or code are needed to access or manipulate shared scientific data, and if so, provide the name(s) of the needed tool(s) and software and specify how they can be accessed.

Specialized tools and software are needed to access or manipulate shared scientific data, including REDCap for electronic data capture, specific image analysis software for Doppler studies and 3D fetal thigh images (e.g., DICOM viewers and analysis tools), and statistical software packages (e.g., R, Python libraries) for data analysis. These tools can be accessed through institutional licenses or open-source repositories.

Element 3: Standards:

State what common data standards will be applied to the scientific data and associated metadata to enable interoperability of datasets and resources and provide the name(s) of the data standards that will be applied and describe how these data standards will be applied to the scientific data generated by the research proposed in this project. If applicable, indicate that no consensus standards exist.

Common data standards will be applied to the scientific data and associated metadata to enable interoperability of datasets and resources. This includes adherence to standardized protocols for data collection (e.g., CDISC for clinical data), imaging data formats (e.g., DICOM for medical images), and metadata standards (e.g., DataCite for DOI assignment). These standards will facilitate the integration of our dataset with existing research databases and promote collaborative research efforts.

Element 4: Data Preservation, Access, and Associated Timelines:

A. Repository where scientific data and metadata will be archived:

Provide the name of the repository(ies) where scientific data and metadata arising from the project will be archived.

Scientific data arising from this project will be archived in a reputable, publicly accessible repository such as the National Institutes of Health (NIH)'s National Library of Medicine (NLM) databases or other discipline-specific data repositories recommended by the NIH.

B. How scientific data will be findable and identifiable:

Describe how the scientific data will be findable and identifiable, i.e., via a persistent unique identifier or other standard indexing tools.

Data will be made findable and identifiable through the assignment of a persistent unique identifier (e.g., DOI) and indexing in relevant databases and search engines, facilitating easy discovery and access by the research community.

C. When and how long the scientific data will be made available:

Describe when the scientific data will be made available to other users (i.e., no later than the time of an associated publication or end of the performance period, whichever comes first) and for how long data will be available.

The scientific data will be made available no later than the time of publication of the main findings or at the end of the performance period, whichever comes first, and will remain accessible for a minimum of 10 years after the study's completion to allow for extended research and validation studies.

Element 5: Access, Distribution, or Reuse Considerations:

A. Factors affecting subsequent access, distribution, or reuse of scientific data:

NIH expects that in drafting Plans, researchers maximize the appropriate sharing of scientific data. Describe and justify any applicable factors or data use limitations affecting subsequent access, distribution, or reuse of scientific data related to informed consent, privacy and confidentiality protections, and any other considerations that may limit the extent of data sharing.

Access to certain data may be limited due to considerations related to informed consent, privacy, and confidentiality protections for human research participants. However, all efforts will be made to maximize sharing while protecting these interests.

B. Whether access to scientific data will be controlled:

State whether access to the scientific data will be controlled (i.e., made available by a data repository only after approval).

Access to the scientific data will generally not be controlled, with the aim of making data openly accessible through public repositories. However, in cases where privacy or confidentiality concerns necessitate restricted access, a controlled access mechanism may be implemented.

C. Protections for privacy, rights, and confidentiality of human research participants:

If generating scientific data derived from humans, describe how the privacy, rights, and confidentiality of human research participants will be protected (e.g., through de-identification, Certificates of Confidentiality, and other protective measures).

To protect the privacy, rights, and confidentiality of human research participants, all shared data will be de-identified, and additional protective measures such as Certificates of Confidentiality will be used where applicable. Informed consent forms will include information about data sharing to ensure participants are aware of how their data may be used.

Element 6: Oversight of Data Management and Sharing:

Describe how compliance with this Plan will be monitored and managed, frequency of oversight, and by whom at your institution (e.g., titles, roles).

Compliance with this Data Management and Sharing Plan will be monitored and managed by the study's principal investigator, in collaboration with the institution's research integrity office and data management team. Regular reviews (at least annually) will ensure adherence to the plan, address any issues related to data sharing, and update the plan as necessary to reflect changes in policies or project requirements.