# Response to Clarification Points for Data Sharing Agreement

## Definition of Limited Data:

***Original Clause RP1 :*** *The Data Provider will transfer a Limited Data Set to the Data Recipient. A “Limited Data Set” consists of health information that has had all direct identifiers concerning the subject of the record (and his or her employer, family, and household members) deleted; that is, the information excludes all of the following: names; street addresses (excluding suburb, small area or town); telephone numbers; fax numbers; electronic mail addresses; government insurance numbers; medical record numbers; health plan beneficiary numbers; account numbers; certificate/license numbers; vehicle identifiers and serial numbers, including license plate numbers; device identifiers and serial numbers; web universal resource locators (URLs); internet protocol (IP) address numbers; biometric identifiers, including finger and voice prints; and full-face photographic images and any comparable images.*

***Suggested Revision for RP1:*** *The Data Provider will transfer a Limited Data Set to the Data Recipient. A “Limited Data Set” consists of de-identified health information, including essential variables such as unique IDs (study ID and participant ID), date of delivery or date of follow-up for maternal outcomes, and location information at a minimum level of city or town. The Limited Data Set excludes all direct identifiers such as names, street addresses, telephone numbers, email addresses, government insurance numbers, medical record numbers, and other identifying information as specified in the original clause. The purpose of the Limited Data Set is to ensure the privacy and confidentiality of the data subjects while allowing for the analysis of maternal and child health outcomes as described in the Annexes of this agreement.*

## Definition of Study Data:

**Relevant Text:** "Study Data" shall mean data and results produced in the execution of the Study.

**Suggestion:** to include this clause:

**Clause [Number]: Definition and Use of Study Data**

1. **Definition of "Study Data":** "Study Data" shall refer to all data and results produced in the execution of the Study, including:
   * **Source Data:** Original data collected during the study, including all forms of raw data.
   * **Internally Shared Data:** Data that includes limited identifiers, shared internally for the purposes of data validation and analysis.
   * **Completely Deidentified Data:** Data from which all personal identifiers, including but not limited to names, addresses, and contact information, have been removed, ensuring no link back to the individual subjects.
   * **Aggregated Tables:** Data compiled from individual data points and presented in summary form.
2. **Harmonization and Deidentification Standards:** In line with NIH grant funding policies and to facilitate broader research collaborations, all "Study Data" intended for sharing outside the immediate research team shall be harmonised and wholly de-identified. These data will exclude specific geolocations, dates, or any other detail that could potentially re-identify the subjects, including but not limited to the country of origin.
3. **Usage of Harmonized and Deidentified Data:** Harmonized and deidentified data will be made available for further research under controlled access following approval of a data request by the DS-I Africa entity or the HE2AT Center Steering Committee. This ensures all shared data adheres to ethical research standards and protects participant confidentiality.
4. **Compliance and Oversight:** The HE2AT Center Steering Committee will monitor compliance with this data sharing policy and handle requests for data access. The Committee will ensure that data sharing aligns with the goals of the HE2AT Project and the stipulations of this Agreement.
5. **Publication and Data Use:** Publications using "Study Data" must adhere to the HE2AT Center authorship policy and ensure no identifiable data is published. The publication data must be based on the harmonised and deidentified datasets as stipulated in this Agreement.

## Duration of Data Retention:

* **Original vs. Proposed Amendments**

**Clause 2.1: Agreement Duration and Termination**

* Original: "This Agreement shall commence on the Commencement Date and will terminate on 30 June 2026 or upon completion of the Project, whichever event occurs first."
* **Proposed: No changes required.**

**Clause 2.2: Termination Procedures**

* **Original: "**Notwithstanding the abovementioned, either Party may cancel this Agreement with 30 (thirty) days’ prior written notice. On termination of this Agreement, the Data Recipient will immediately discontinue use of the Data and will return all copies of the same to the Data Provider or alternatively, and on the Data Provider’s written instruction, destroy all copies of the Data."
* **Proposed:** "Notwithstanding the above, either Party may cancel this Agreement with 30 (thirty) days’ prior written notice. Upon termination of this Agreement, the Data Recipient shall immediately discontinue use of all Data and shall return or destroy all copies of the Data in accordance with the Data Provider’s instructions, unless otherwise agreed upon in writing."

**Clause 2.14: Data Retention Period**

* **Original:** "The Data Recipient will retain a copy of the Data [internal cleaned] for a period of 5 years after the termination of the over-arching NIH grant agreement (current Project End Date 30 June 2026) for the purposes of concluding and correcting any analysis and publications resulting from the [internal cleaned] Data."
* **Proposed:** "The Data Recipient will retain a copy of Completely Deidentified Data and Aggregated Data indefinitely to support ongoing and future research under the terms agreed upon for open access data. However, Source Data and Internally Shared Data will be retained only until the termination of the over-arching NIH grant agreement (current Project End Date 30 June 2026), to conclude and correct any analysis and publications resulting from the Project. Any retention of Source and Internally Shared Data beyond this period must be expressly approved by the Data Provider and in compliance with applicable regulations."

**Clause 2.15: Authority and Consent for Data Transfer and Use**

* **Original:** "By signing this Agreement, the Data Provider confirms that it has the authority to transfer the Data and consent to provide the Data to the Recipient for use for the duration of this Agreement and as provided for in Clause 2.14."
* **Proposed:** "By signing this Agreement, the Data Provider confirms that it has the authority to transfer the Data and consents to provide the Data to the Recipient for use in accordance with the terms outlined in this Agreement and particularly as provided for in Clause 2.14 regarding data retention and use post-project completion."

**Additional Clause: Data Ownership and Transformation**

* **New Clause Suggested:** "Data Ownership and Transformation: Following the deidentification and harmonisation process, the ownership of the Completely Deidentified Data and Aggregated Data transfers to the HE2AT Center, allowing for indefinite retention and use within the terms set for open access data. This transformation ensures the utility of the data for ongoing and future scientific research while adhering to ethical standards and privacy laws."

## Inclusion of CHEAQI in HE2AT Center:

**Relevant text:**

* *Clause 2.7:* Data Recipient will use the Data only for purposes of the Project. If the Data Recipient seeks to use Data for other purposes, the Data Recipient will obtain written consent from Data Provider, either by an amendment to this Agreement or a new agreement, before such use.

**Suggestion:**

To ensure a smooth and efficient process when new partners are added to the HE2AT Center, it's important to include provisions in the generic Data Transfer Agreement (DTA) that specifically address the integration of new collaborators. Here’s a proposal for such a clause designed to facilitate easy incorporation of new partners without requiring extensive renegotiations or amendments.

Clause [X]: Incorporation of New Partners

**Proposed Text:**

"In order to adapt to the evolving nature of research and collaboration within the HE2AT Center, this Agreement anticipates the addition of new partners and collaborators as the project progresses. The following provisions are established to streamline the integration of new partners into existing agreements:

**Notification and Approval:**

The Data Recipient will notify the Data Provider in writing of the intent to add a new partner or collaborator to the HE2AT Center. Approval from the Data Provider must be obtained before the new partner is formal engaged. Such approval will not be unreasonably withheld, provided the new partner agrees to comply with the terms set forth in this and all related agreements.

**Standardised Accession Protocol:**

Upon approval, new partners will be integrated through a standardised accession protocol, which includes signing an Accession Agreement that binds them to the terms and conditions of the existing DTA. This Accession Agreement will outline the rights and responsibilities of the new partner, ensuring alignment with the consortium’s objectives and compliance with data protection standards.

**Modification of Agreement:**

This Agreement includes a provision for modification to accommodate new partners' specific needs and contributions. Such modifications will be appended to the existing Agreement in the form of Schedules, which detail the nature of data to be shared, specific use cases, and any additional confidentiality and security measures required.

**Seamless Integration:**

The process for integrating new partners will be designed to ensure seamless collaboration and data sharing within the HE2AT Center. The consortium will maintain a flexible yet secure framework that facilitates the inclusion of new data, tools, or methodologies introduced by new partners.

By establishing these provisions, the HE2AT Center aims to foster continuous collaboration and growth, ensuring that new partners can be efficiently incorporated into ongoing projects with minimal disruption and maximum compliance with governing data sharing policies and regulations."

## Permissions Required by Data Providers:

*Relevant Clauses:*

* *Clause 2.5: Data Provider retains ownership of the Data and retains all rights to distribute the Data to other third parties. Data Provider warrants its authority and that it has obtained the necessary consent required to provide the Data to the Data Recipient.*
* *Clause 2.10: The Parties acknowledge their obligation(s) to comply with Data Protection Legislation and that violation of the Data Protection Legislation may subject them to applicable legal penalties.*
* *Clause 2.15: By signing this Agreement, the Data Provider confirms that it has the authority to transfer the Data and consent to provide the Data to the Recipient for use for the duration of this Agreement and as provided for in Clause 2.14.*
* *Clause 4: Each Party will comply with Data Protection Legislation in relation to the performance of its obligations under this Agreement.*

***Confirmation****:* The current text adequately covers the ethical and legal responsibilities of the data provider, with no need for alterations.

*For open access data* definitions, use ontology for describing the data

## Transfer of Data from SA to Zim/UPGC:

*Relevant Clauses:*

* *Clause 7: In the event that it is necessary for the Data Recipient to transfer Personal Data across national borders to authorised Collaborator/s or other authorized third parties (as may be agreed between the Parties), the Party providing the Data will ensure the lawful export of the Personal Data and shall enter into a separate agreement governing such transfer on terms no less stringent than the terms set out herein.*
* *Clause 7.2: In the event that the Data is transferred to a jurisdiction where POPIA does not apply, the respective Party transferring the Data undertakes that the Data will only be transferred to a jurisdiction with adequate protection as set out under Section 72 (1) of POPIA.*

*Suggestion: Can we drop these clauses?*

*Question for Donrich in the ELSI: How do these clauses apply to the transfer of data from South Africa to Zimbabwe/UPGC, especially in cross-border data transfers and compliance with POPIA and other relevant national legislation? Are additional considerations or agreements required to ensure lawful and secure data transfer between these jurisdictions?*

## Data Impact Assessment and Logs:

* *Clause 11:* Records/Data Processing Register

*Clarification from Chris and Lisa:* What is the scope of the data protection impact assessment required under this clause, and are we currently fulfilling this requirement?

Checked and this was described as being the Trello Board

## List of Variables Requested:

* *Suggested Change:*

The below list of variables is indicative and the final variable list to be finalised during discussion between the Data Provider and Recipient, based on data availability and relevance.

Individual participant data for a limited set of variables from the original dataset/s relating to maternal and child health, which may include: