Research Data Governance and Sharing Framework

Updates

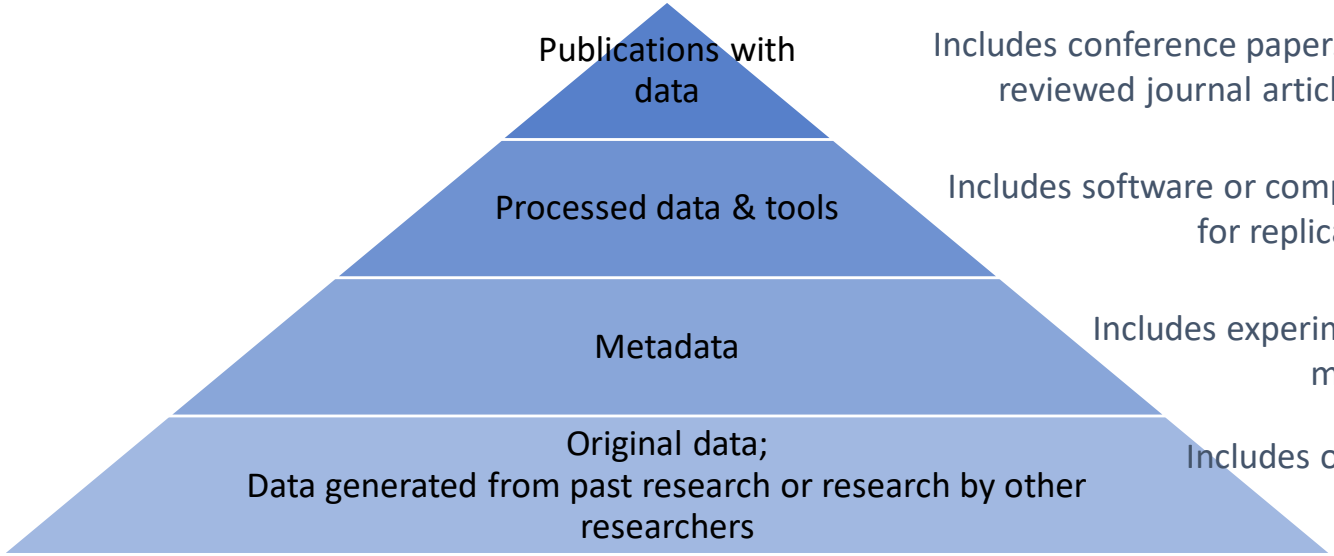
Framework for Research Data Governance and Sharing

- NMRC developed a Framework for Research Data Governance and Sharing, which was endorsed by the HBMS EXCO in October 2014.
- The framework requires publicly-funded project to:
  a) Allow open access to peer-reviewed publication
  b) Share final research data for projects seeking funding of S$250,000 and above
- The Research Data Governance and Sharing Working Committee recommended a 2-phase implementation of the framework.

Phase 1	Phase 2
• Requires open access to peer-reviewed publication • Applies to all existing or new projects w.e.f. **May 2015**	• Requires sharing of final research data • Applies to new projects seeking funding of $250,000 and above

Final Research Data

- Final Research Data refers to recorded factual material commonly accepted in the scientific community as necessary to document and support research findings.



Does not include laboratory notebooks, preliminary analyses, drafts of scientific papers, plans for future research, peer review reports, communications with colleagues, or physical objects, such as gels or laboratory specimens

Broad Principles for Research Data Sharing

Timeliness

- Data to be made available at the earliest feasible opportunity.
- Delays or restrictions on data sharing may be justified to gain intellectual property protection or to further develop technology for public benefit.
- Peer-reviewed publications to be made publicly available no longer than 12 months after official date of publication.

Human Subjects and Privacy

- Rights and privacy of human subjects to be protected at all times.
- Data shared to be de-identified.
- Informed consent to include explicit provision for sharing of data for wider or future research use, subject's approval to be contacted should study show that clinical intervention is required, and consent to collect clinical data from national-level databases.

Governance of Data Access

- Data Access Oversight Committee (DAOC) to oversee the implementation of policy, resolve disagreements relating to data access requests and handle appeals.

Supplemental Baseline Requirements

Feedback from institutions for NMRC to put in place high-level baseline requirements to assist institutions in formulating the institutional policies to support the sharing of research data.

Baseline requirements for sharing of research data in 6 areas were developed:

1. Type of Research Data to Share
2. Timeframe for Sharing Research Data
3. Where to Store Research Data
4. Human Subjects and Privacy
5. Access of Data and Governance
6. Data Retention Period

Supplemental Baseline Requirements – Summary (1)

Type of Research Data to Share	• Final research data* and its associated metadata from all type of studies. • For **clinical data**, should there be reasonable limitation to sharing, a **minimum set of clinical data, including basic demographic (age, gender and race), and disease or condition (diagnosis)**, should be included.
Timeframe for Sharing Research Data	• Data to be shared **no longer than 12 months after official date of publication** if dataset is used for publications. Otherwise, data should be shared within **24 months after grant completion.**
Where to Store Research Data	• A copy of the final research data to be stored in **a centralized research data repository** developed by IHIS.

* Final research data refers to recorded factual material commonly accepted in the scientific community as necessary to document and support research findings. This does not mean summary statistics or tables; rather, it means the data on which summary statistics and tables are based.
http://grants.nih.gov/grants/policy/data_sharing/data_sharing_guidance.htm#fin

Supplemental Baseline Requirements – Summary (2)

Data Retention Period

- Research Data residing in the centralized data repository will be retained for a **minimum period of 10 years** and will be subject to NMRC Data Repository Archival protocol.

Human Subjects and Privacy

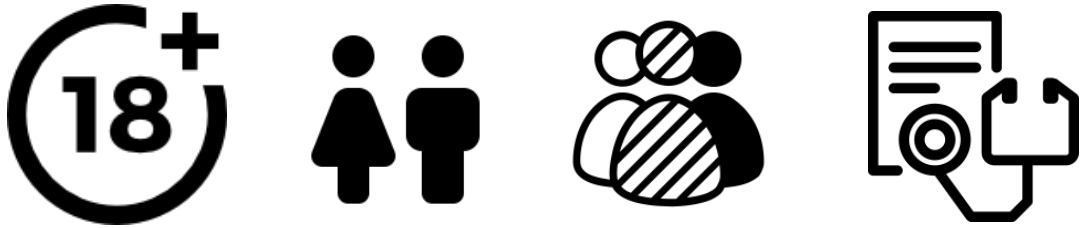
- Institutions to provide Institutional Certification prior to fund disbursement to assure that submission of data is consistent with the Data Sharing policy, and that **an IRB has reviewed the investigator's data sharing plan**.
- Investigators will submit **de-identified data** (using common de-identification standards) to the NMRC Data Repository.
- **Controlled access** for all human subject related datasets.

Access of Data and Governance

- For managed access data, qualified investigators are to submit a **Data Access Request (DAR)** endorsed by the requesting Investigator's institution. The endorsed DAR will be **reviewed by the data contributor's institution-based Data Access Committee/ Representative**.

1. Type of Research Data to Share

- Researchers are required to share data for all types of studies, including final research data and its associated metadata.
- With respect to clinical data, if there are reasonable limitations to sharing, a minimum set of clinical data should be shared.



Age **Gender** **Race** **Diagnosis**

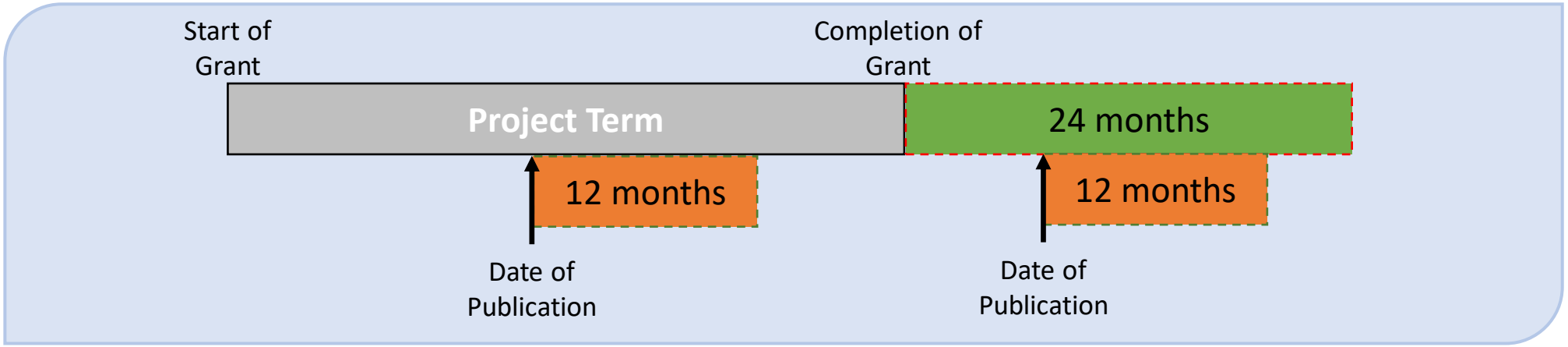
2. Timeframe for Sharing Research Data

Dataset used in publication

No later than 12 months after official date of publication

Dataset not used in publication

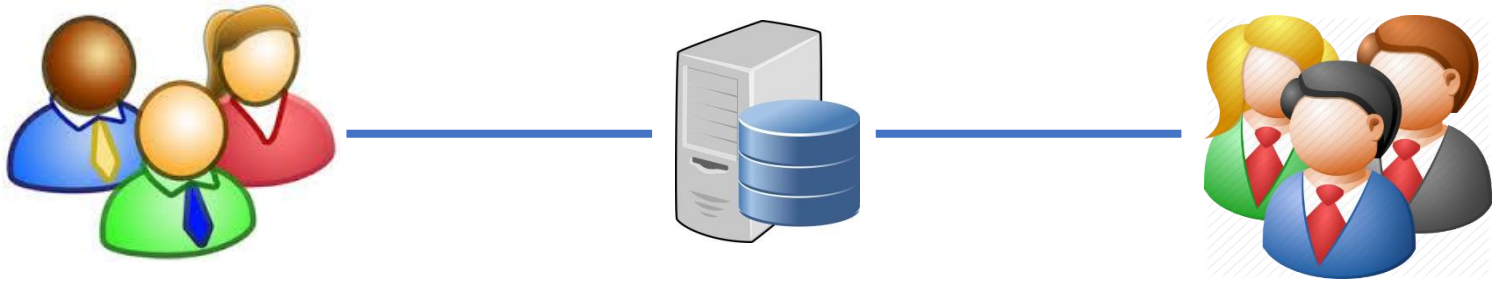
No later than 24 months after grant completion



Extension of deadline may be sought if there are strong justifications for the extension (e.g. manuscript under review) and evidence of the same, subject to NMRC's approval.

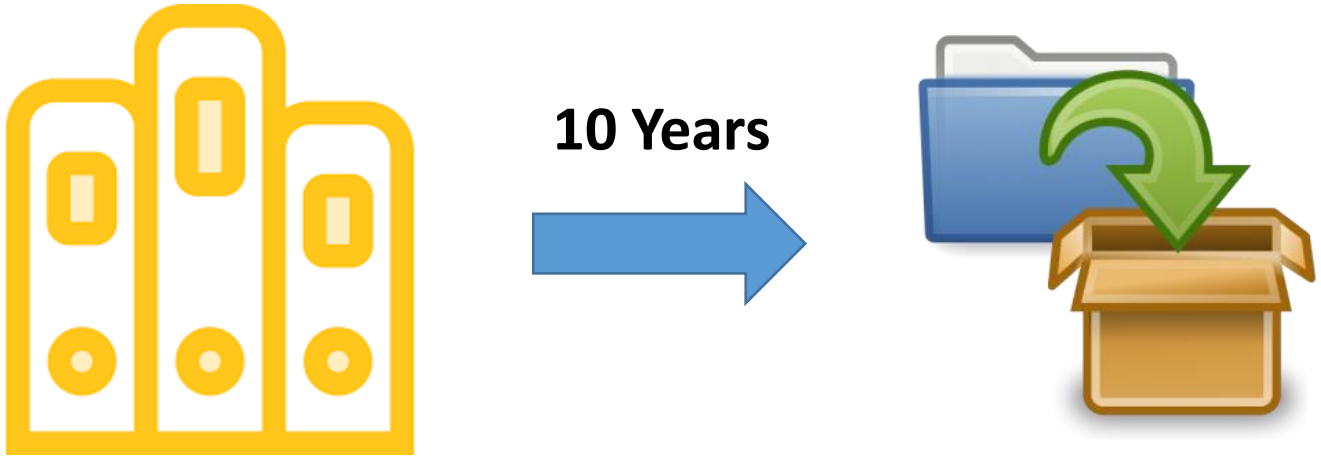
3. Where to Store Research Data

A copy of the research data to be stored in a centralised research data repository developed by IHiS.



4. Data Retention Period

- Research Data residing in the Research Data Repository will be retained for a minimum period of ten (10) years.
- NMRC will work with IHiS to develop plan for long-term/permanent archival of research data.



5. Human Subjects and Privacy

Review by IRB

Data Sharing Plan requires review and endorsement by IRB before the awarded project may commence

Controlled Access

Controlled access for all human subject related datasets

De-identified Data

Investigators are to submit de-identified data (using common de-identification standards) to the Research Data Repository

Linkage of Data

Linking of data (if required) will be performed by Trusted Third Party and can be done without data re-identification when the data involved use the common de-identification standards

5. Human Subjects and Privacy

Review by IRB

Data Sharing Plan requires review and endorsement by IRB before the awarded project may commence

- The IRB review should include <u>assessment/consideration</u> on
  - Consistency of data sharing plan with the informed consent procedures;
  - Adequacy of de-identification standards to protect human subject involved;
  - Sensitive data that may potentially lead to stigmatization;
  - Appropriate parameters for use of data through the use of data use limitations that could minimise the potential for harm.
- Investigators and their institutions may request removal of data on individual participants from the repository in the event that consent is withdrawn. However, data that have already been distributed to approved users for research will not be able to be retrieved.

6. Access of Data and Governance

Open Access

Project Information

Final Research Data Metadata

Available to public (without login)

Controlled Access

Final Research Data (De-Identified)

Available to specific researchers with approved **Data Access Request**.

For Controlled Access Datasets

Qualified investigators* are to submit a Data Access Request (DAR) that includes a proposed Research Use Statement^ and a signed Data Use Certification Agreement#.
The DAR must be endorsed by the requesting Investigator's institution.

The endorsed DAR will be reviewed by the Data Contributor's institution-based Data Access Committee/ Representative.
Decisions to grant access should be made based on whether the proposed use of the dataset is scientifically and ethically appropriate, and whether the request conforms to the Data Sharing Policy and is consistent with the Data Use Limitations@ prescribed for the dataset.
Appeals for rejected request can be evaluated by the NMRC Data Access Oversight Committee (DAOC).

6. Access of Data and Governance

Footnote

* **Qualified Investigators** must be employees with a primary appointment in a local public institution at a level equivalent to a Principal Investigator. Laboratory staff and trainees such as graduate students, and postdoctoral fellows are not permitted to submit data requests.

^ **Research Use Statement** (cap at 2 pages with Arial font 10) should include the following components:

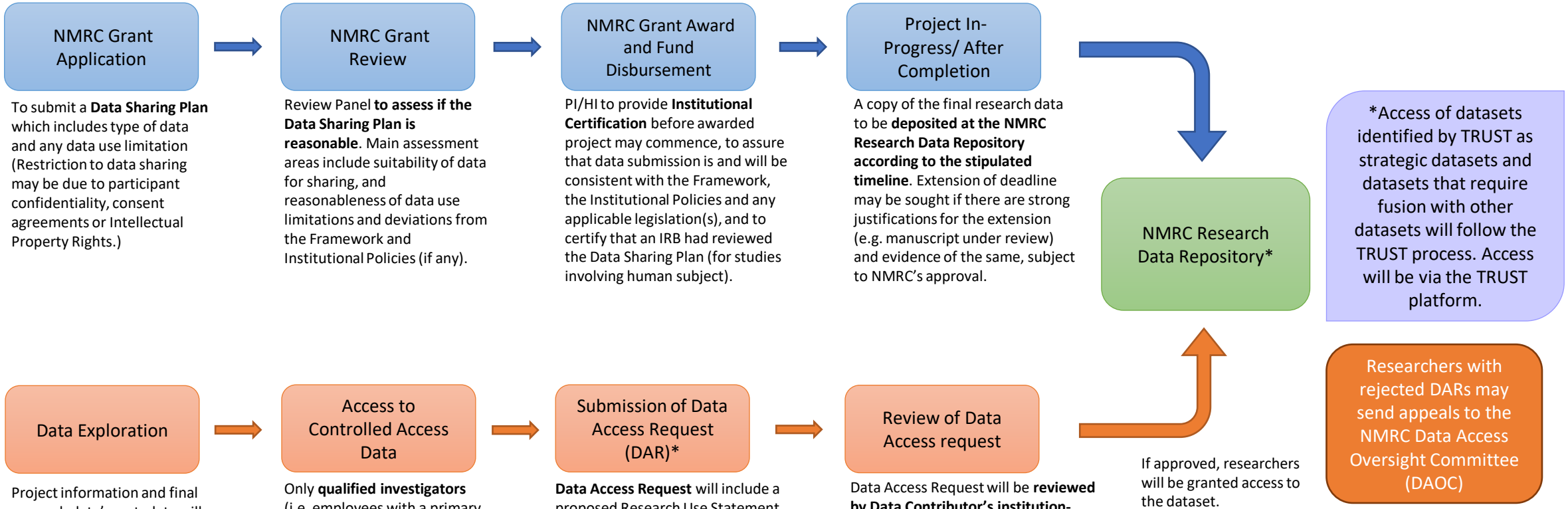
- Objectives of the proposed research;
- Study design and analysis plan;
- Explanation of how the proposed research is consistent with the data use limitations of the requested dataset; and
- Brief description of any planned collaboration with researchers at other institutions, including the name of the collaborator(s) and their institution(s).

Data Use Certification Agreement would include the following terms:

- To use the data only for the approved research;
- To protect the confidentiality of the data;
- To not attempt to identify or contact individual participants from whom the data were obtained;
- To not distribute the data to any third-party entity or individual;
- To adopt appropriate data security measures to ensure that only authorized individuals can gain access to the data files;
- To follow applicable law, regulations and local institutional policies and procedures for handling data
- To acknowledge in all oral or written presentations, disclosures, or publications the contributing investigator(s) who conducted the original study, the funding organization(s) that supported the work and the specific dataset(s); and
- To agree to the listing of requester name and institutional affiliation, and the approved use of dataset.

@ **Data Use Limitations** will be part of the Data Sharing Plan submitted by data owners during grant application. The limitations should be based on the terms of informed consent of the study participants from whom the data are generated, or based on valid justifications. Example of limitations include dataset can only be used for nonprofit organization, or dataset can only be used for health/biomedical research, or dataset can only be used for research related to a specific disease.

Research Data Sharing Flowchart



What is expected from Researcher (or Data Contributor) of NMRC-Funded Research

Data generated from NMRC-funded research could potentially be identified as strategic dataset by TRUST.

	Datasets identified by TRUST (e.g. OF-LCG project)	**Datasets not identified by TRUST**
Sharing Pre-agreement	Establish contributor pre-agreement with TRUST to establish terms of sharing	NA
Cleaning and data curation	Data contributor is responsible for preparing de-identified final research data and meta-data for submission. TRUST data cleaning team will work with data contributor to share data standards, mapping tables, algorithms etc. to support the data cleaning process.	Data contributor is responsible for preparing de-identified final research data and meta-data for submission.
Storage in RDS	Adhere to the timeframe* to deposit the final research data in RDS. The cleaned de-identified data from TRUST will be accepted as final research data deposited to fulfil NMRC's requirement.	Adhere to the timeframe* to deposit the de-identified final research data in RDS.

** As per the timeframe defined under the supplemental baseline requirement, i.e. no longer than 12 months after official date of publication if dataset is used for publications. Otherwise, within 24 months after grant completion.*

What is expected from Researcher (or Data Contributor) of NMRC-Funded Research

	Datasets identified by TRUST (e.g. OF-LCG project)	**Datasets not identified by TRUST**
Access	Access request will follow TRUST process, involving TRUST DAC comprising institutional rep (i.e. PHIs, IHLs) Upon approval by TRUST DAC, Data would be drawn directly from agreed source system under the pre-agreement into TRUST. Access of fused research data, govt admin data, real world data etc. is via TRUST platform.	Access request will be assessed by institution-based DAC/Rep. For access to such dataset on its own (i.e. no fusion with other datasets), BRAIN will support the use case. When fusion with other datasets (e.g. other research data, govt admin data, real world data) is required, researchers are to submit request to TRUST.

Clarifications

NMRC Research Data Repository, BRAIN and TRUST

Researchers deposit a copy of the research data (in de-identified form) in the NMRC Research Data Repository

NMRC Research Data Repository

Requested data are transferred to TRUST/BRAIN for data requesters' access upon institution's approval. Original data in repository remains intact.

TRUST Platform

- Not a data repository
- Serves as a data access/analytical platform for data requester who requires linkages with other data
- Linking of data performed by Trusted Third Party. This can be done without data re-identification due to the common de-identification standards among major platforms.
- Strategic research data identified by TRUST available for wider research community

BRAIN Platform

- Not a data repository
- Serves as data access/analytical platform for data requester who does not require linkages with other data

Requested data will only be made available to data requesters via BRAIN/TRUST in de-identified form

Note: The access to requested data via other platforms would not be possible at the moment as various aspects including data security concerns would need to be looked at.

Research Data to Deposit

- The general guiding principle to what constitute final research data **would be what researchers themselves find sufficient to validate research findings and support future research work.**
- Data standards and formats are not stipulated at this stage.
- Sharing of partial data is possible, if restrictions are deemed reasonable.
  - Such limitations (e.g. omission of certain data, datasets only comprise data from 60 out of 80 subjects, or only anonymized data is collected) should be made known in the meta data so that data requesters/public are aware.

Data Owners' Responsibilities and Rights

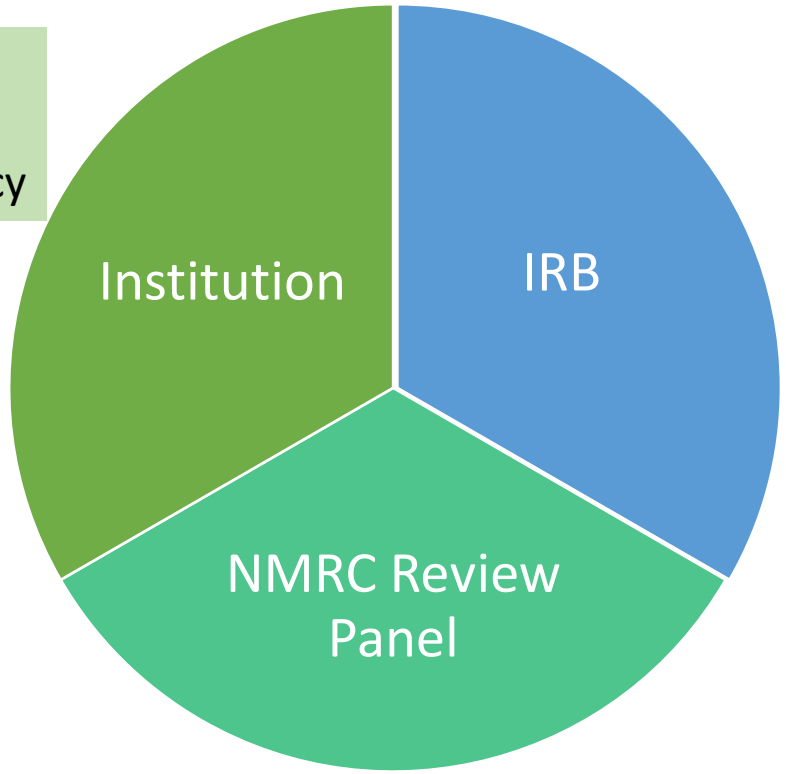
- Data contributor/data contributing institution, who are the data owners, will be responsible for the integrity, quality and accuracy of the final research data submitted.
- PIs/HIs continue to have the rights to use their research data for future research and for sharing based on their discretion.
- Although unpublished data should be shared within 24 months after grant completion, request to delay the stipulated timeline may be sought if there are strong justifications (e.g. manuscript under review, intellectual property protection) and evidence of the same, subject to NMRC's approval.

Informed Consent for Data Sharing

- Informed consent should be sought with option to opt out
  - Prospective participants should understand what is expected to happen with their data, and the data shared (if any) will be in de-identified form.
- For past studies where informed consent for sharing of data for wider/future research use was not included, this could be stated in the data sharing plan for consideration of data sharing to be waived.
- Linkage consent would facilitate linkages to other datasets in the future, but requesters may seek IRB approval for waiver of consent downstream.

Responsibility in Reviewing Data Sharing Plan – IRB, Institution, NMRC Review Panel (1)

- Carry out administrative checks to ensure compliance and consistency



- Assess risks and check that human subjects are sufficiently protected in relation to data sharing.
- Ensure that necessary steps/limitations are in place to adequately protect the human subject involved.
- E.g. flag up if there are inconsistencies between the informed consent and the data use limitations proposed by investigators.

- The IRB's decision will take priority over the assessment of the NMRC Review Panel with regard to human subject protection.

- Assess the reasonableness of the data sharing and data use limitations (e.g. is it justifiable to limit sharing due to intellectual property rights protection, is it reasonable to limit data use to certain research areas)

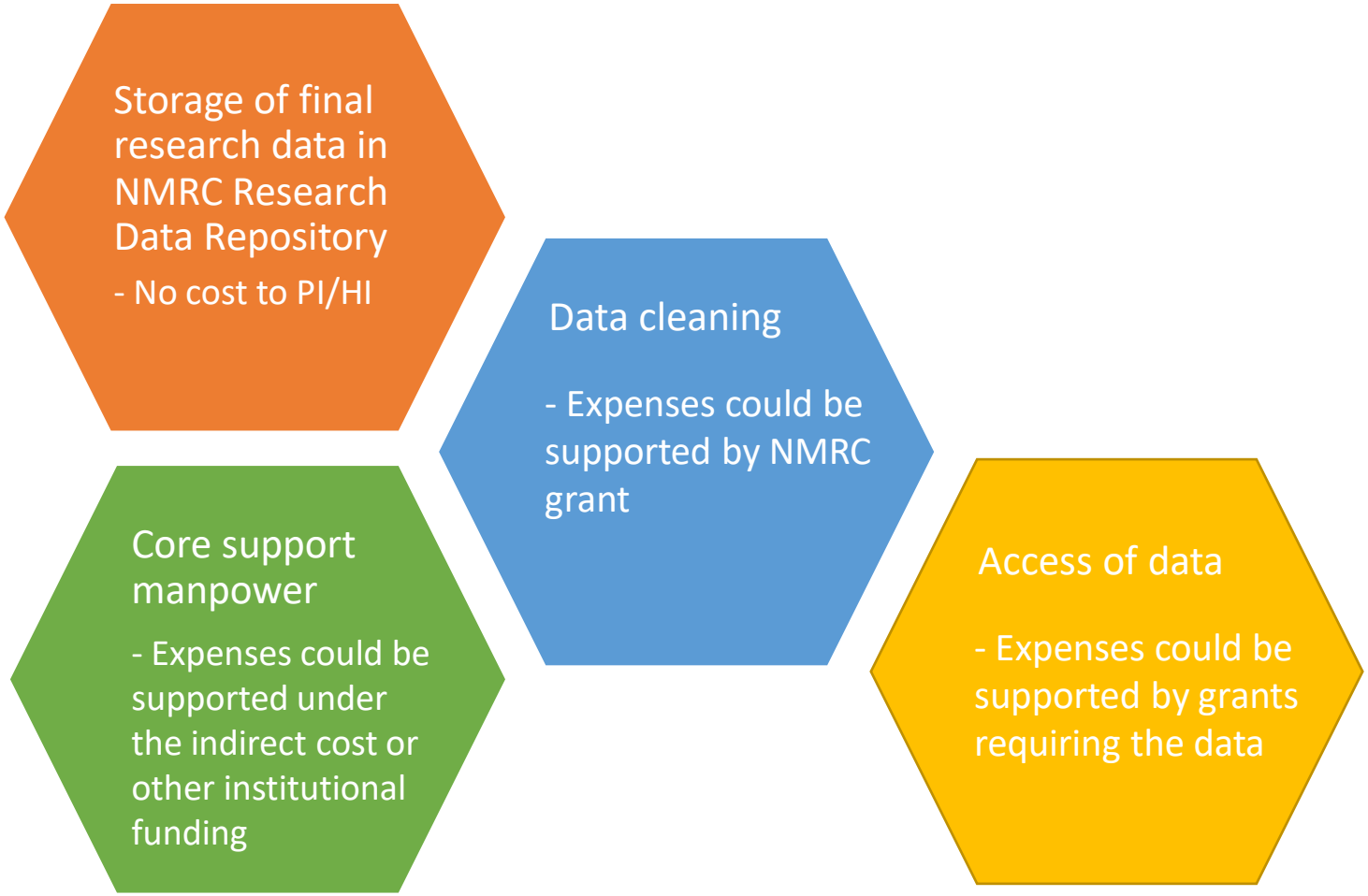
Responsibility in Reviewing Data Sharing Plan – IRB, Institution, NMRC Review Panel (2)

The de-identification standards should be largely in line with the Framework on Anonymisation Standards for the Public Healthcare Sector (MOH circular sent to MOHH group entities in Mar 2022).

ANNEX A – DATA TREATMENT CONTROLS FOR IDENTIFIERS

S/N	I. Identifier	II. Example(s) include, but not limited to	III. MOH Framework on Anonymisation for the Public Healthcare Sector – Data Treatment Controls for Identifiers
1	**Name**		To remove full and partial names.
2	**Address**	Residential address, block number, unit number, street name.	To remove address.
3	**Zip code, postal code**		To retain initial 3-digits of postal code and similar geographical divisions/sub-divisions for all properties.
4	**Date of Birth**		To retain only month and year.
5	**Contact number**	Telephone number, fax number, mobile number.	To remove.
6	**URLs, email address, IP address**		To remove.
7	**Vehicle identifiers and serial numbers**	License plate number.	To remove.
8	**Personal identification**	NRIC number, FIN number, passport number	

Cost Related to Research Data Sharing



Data Access Request Assessment

Data Request Criteria

- Data requesters should be employees with a primary appointment in a local public institution at a level equivalent to a Principal Investigator
- Data access requests are investigator-specific and with specific use case.

NMRC Check

- NMRC will check the data requestor's eligibility before sending the requests to the relevant data contributing institution for assessment.

Assessment

- The institution/cluster-based Data Access Committee (DAC) will assess any data access requests involving datasets contributed by the PI/HI.

Appeals

- The NMRC Data Access Oversight Committee will serve to mediate between the requester and the data contributing institution where necessary.

Approach for Collaborative Projects

- Data might be owned by different data contributors/institutions in a collaborative projects.
- The project team is encouraged to explore feasible arrangements to facilitate requests to access the research data.
- For example, a project-level data access committee comprising representatives of all data contributing parties could be set up for the lead PI's institution's data access committee to consult when granting approval to data requests, or the different parties could defer the decision on the study data to the lead PI's institution's data access committee.

Implementation Timeline

- Institutions need not include specific clauses related to NMRC data sharing requirements in their institutional policies.
- We would appreciate if the institutions could get the institutional policies ready **by end Oct 2022**.