**GUIDANCE ON THE DATA SHARING PLAN TO BE SUBMITTED TOGETHER WITH AN APPLICATION TO THE NATIONAL MEDICAL RESEARCH COUNCIL FOR FUNDING OF RESEARCH (“GUIDANCE”)**

1. **Overview**
	1. All persons: -
2. making an application (“**Application**”) for funding of at least S$250,000 of direct cost[[1]](#footnote-1) from the National Medical Research Council (“**NMRC**”) under the NMRC Research Grant Scheme for research (“**Research**”) listed in Annex A of the Research Data Governance and Sharing Framework (“**Framework**”); and
3. submitting the Application to the NMRC on or after the Effective Date;

shall: -

1. submit a data sharing plan (“**Data Sharing Plan**”) in accordance with the template in Annex A of this Guidance, together with the Application; and
2. if the Application is successful, be required to share data from the research which the funding is for, pursuant to the Framework.
	1. In this document, the “**Effective Date**” means 1 January 2024.
3. **Data Sharing Plan**
	1. The Data Sharing Plan is a key tool to (i) set out the considerations and plans to share research data under the Framework; and (ii) justify requests for additional resources and funding in relation to the sharing of research data under the Framework.
	2. The research personnel to be identified in the Application as the overall lead in the conduct of the research (“**Lead Principal Investigator**”) shall ensure that: -
4. he/she is familiar with the objectives and principles of the Framework; and
5. the Data Sharing Plan is duly completed and submitted together with the Application.
	1. Annex B of this Guidance sets out the explanatory notes for the Data Sharing Plan template.
	2. The Data Sharing Plan will be reviewed and assessed separately from the other components of the Application, by the NMRC’s review panel or equivalent body, in accordance with the relevant guidelines in place from time to time. Upon the review and assessment, the NMRC’s review panel or equivalent body may request for the Data Sharing Plan to be revised. If so, the Lead Principal Investigator and the Host Institution shall ensure that the Data Sharing Plan is revised accordingly and submitted to the NMRC, within the timeframe indicated by the NMRC.

**Annex A**

**DATA SHARING PLAN TEMPLATE**

The Data Sharing Plan shall be submitted in accordance with the template set out below. The notes in italics under each section of the template provide further context and guidance on how the section should be completed. Please refer to Annex B (Explanatory Notes to the Data Sharing Plan Template) for further information.

**1. Description of Research and Research Data**

**1.1 Brief Description of the Research Study** *Briefly describe the nature of the research study (in not more than 3 lines/sentences).*

**1.2 Brief Description of the Research Data and Analysis to be Undertaken**

*Briefly describe the research data that will be collected/generated, and the plans for/scope of analysis to be undertaken.*

**2. Restriction on Data Sharing**

*A copy of the anonymised Final Research Data is expected to be stored in the NMRC Research Data Repository for sharing under the Framework.*

*You should share all the Final Research Data under the Framework within the timeframe stated in the Framework. Should you anticipate that this cannot be done, please state: -*

1. *the scope of the Final Research Data that you will likely not be sharing under the Framework, or likely not be sharing under the Framework within the timeframe stated in the Framework;*
2. *explain why that Final Research Data likely cannot be shared under the Framework or likely cannot be shared under the Framework within the timeframe stated in the Framework. Only strong justifications will be accepted; and*
3. *include a proposal on alternative means in which that Final Research Data can be shared e.g. by anonymising or aggregating that Final Research Data, obtaining the relevant consent to disclose that Final Research Data; obtaining the relevant copyright permissions or sharing that Final Research Data at a later, more appropriate time (in which case, an extension of time should be sought pursuant to the Framework).*

*“****Final Research Data****” means, in relation to the Research, recorded factual material commonly accepted in the scientific community as necessary to document and support research findings, regardless of whether the data is used to support publications. Final Research Data that are clinical in nature shall include the human subject’s basic demographics (including but not limited to the human subject’s age, gender and race), disease and condition (i.e. the diagnosis), unless the information is not collected in the Research. Final Research Data shall 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.*

**Please click here to choose an option**

If there is restriction on data sharing, please provide your justifications in the text box below. Otherwise, please enter ‘NA’.

**3. Data Use Limitation**

*Indicate any anticipated limitations to using the Final Research Data shared under the Framework. Such limitations should be minimised where possible and be substantiated with valid reasons e.g. where the limitation is due to limited consent obtained for the use and disclosure of the Final Research Data. Examples of such limitations include the use of the Final Research Data by non-profit organisations only, the use of the Final Research Data for health and/or biomedical research only, or the use of the Final Research Data for research related to a specific disease only.*

**Please click here to choose an option**

If there is Data Use limitation anticipated, please provide your justifications in the text box below. Otherwise, please enter ‘NA’.

**4. Institutional Review Board (IRB) Endorsement (For a Research Study that involves human subjects)**

*For research study that involves human subjects, please state whether an IRB has reviewed and endorsed the Data Sharing Plan.*

**Please click here to choose an option**

**5. Lead Principal Investigator’s Undertaking**

*The Lead Principal Investigator shall ensure that the Data Sharing Plan is (i) in accordance with the Framework, and the Lead Principal Investigator’s institution’s data sharing/management policy; and (ii) in compliance with all applicable legislation prescribing requirements in relation to data.*

 Please tick the box to undertake the following.

|  |  |
| --- | --- |
| [ ]   | I, the Lead Principal Investigator, shall ensure that the submitted data sharing plan is in accordance with the Research Data Governance and Sharing Framework, and my institution’s data sharing/management policy, and in compliance with all applicable legislation prescribing requirements in relation to data. |

Name and signature of Lead Principal Investigator:

Date:

**SUPPLEMENTARY DATA MANAGEMENT QUESTIONS**

This section serves to provide more insights on the data being collected and how it is managed in the proposed project, in complement to the data sharing plan.

Please note that this section will not be assessed as part of the Data Sharing Plan.

**A1. What data will you collect, create or reuse?**

*Describe format of data (e.g., text, numeric, audio-visual, models, computer code, discipline-specific, instrument-specific) and give an indication of the anticipated volume of data. Outline and explain your choice of format and consider the implications of data format and data volumes in terms of storage, backup and access.*

*Indicate usage of any secondary data. This could include any existing data, data from earlier projects or third-party sources such as administrative data available in TRUST (you may refer to the TRUST webpage for more information on the data available in TRUST –* [*https://trustplatform.sg/collaborate/access-trust-data/*](https://trustplatform.sg/collaborate/access-trust-data/)*).*

**A2. How will the data be collected or created?**

*Describe data collection method (e.g., experimental (generated by lab equipment), computational/simulation (generated from computation models), observational (recordings of specific phenomena at a specific time or location), derived (produced via processing or combining other data), reference (extracted from published and/or curated datasets)).*

*Specify any community agreed or other formal data standards used (with URL references) to enable interoperability.*

*Explain how the consistency and quality of data collection will be controlled and documented. This may include processes such as calibration, repeat samples or measurements, standardised data capture or recording, data entry validation, peer review of data or representation with controlled vocabularies.*

**A3. How will collected data be linked?**

*Describe how utility of collected data can be further leveraged with longitudinal data (past and present) from administrative data available in TRUST or other sources to support the project, if applicable. Applicants are encouraged to plan ahead for secondary use of datasets and data linkages (e.g., obtaining necessary consent to link data with administrative data available in TRUST or other sources and secondary use of datasets) to expand the utility of the data collected.*

**A4. How will the data be stored and backed up during the research?**

*Describe where and how you will store the data during the research.*

*Describe the backup and archiving regime you will use to back up all your data to prevent its loss.*

**Annex B**

**EXPLANATORY NOTES TO THE DATA SHARING PLAN TEMPLATE**

*NB: The notes in italics under each section of the table below provide further context and guidance on how the section should be completed.*

| **Sections to be completed in the Data Sharing Plan** | **Remarks** |
| --- | --- |
| 1. **Description of Research and Research Data**
 | In order to provide context to the Data Sharing Plan, you will need to briefly describe the research study and the research data that will be collected and/or generated. |
| * 1. **Brief Description of the Research Study**

*Briefly describe the nature of the research study (in not more than 3 lines/sentences).* | * You may use phrases such as the following to describe the nature of the research study: exploratory study, laboratory study, animal study, human study, randomised controlled trial, clinical trial, cohort study, longitudinal study, prospective/retrospective study, observational study, case-control study.
* *Example A: This is a prospective observational study of 400 smokers and 400 non-smokers, to investigate the correlation between smoking and lung cancer.*
* *Example B: This is a single arm phase II clinical trial to evaluate a novel drug for liver cancer. 50 patients will be recruited and will be followed up for a period of 24 months.*
 |
| * 1. **Brief Description of the Research Data and Analysis to be Undertaken**

*Briefly describe the research data that will be collected/generated, and the plans for/scope of analysis to be undertaken.* | * You may use phrases such as the following to describe the type of research data that will be collected and/or generated: quantitative/qualitative data, data generated from surveys, clinical measurements, interviews, medical records, electronic health records, administrative records, genotypic data, images, tissue samples.
* If the research data includes unique data that cannot be readily duplicated, this should be highlighted here (e.g., large surveys that are expensive to replicate, studies on unique population, studies conducted at unique times).
* *Example A: Medical records and patient-reported questionnaire on lifestyle and dietary intake will be collected. Clinical tests including blood and urine tests will be carried out at baseline and follow-up assessments at 3 time points: 6 months, 12 months and 24 months. Correlation analysis will be undertaken at the end of second year. The study will involve pregnant human subjects with rare genetic diseases, and it is unlikely that the data can be readily duplicated.*
* *Example B: Patient demographics, baseline disease characteristics, and therapy outcomes including recurrence, stage progression and survival will be collected. Biochemical characteristics, including blood counts, liver function tests and renal function tests, for baseline and subsequent follow-up assessments will be measured and recorded. Tumour assessments using magnetic resonance imaging will be conducted every 4 weeks. Pharmacokinetic analysis and statistical analysis will be carried out.*
 |
| 1. **Restriction on Data Sharing**
 |  |
| *A copy of the anonymised Final Research Data is expected to be stored in the NMRC Research Data Repository for sharing under the Framework.**You should share all the Final Research Data under the Framework within the timeframe stated in the Framework. Should you anticipate that this cannot be done, please state: -* 1. *the scope of the Final Research Data that you will likely not be sharing under the Framework, or likely not be sharing under the Framework within the timeframe stated in the Framework;*
2. *explain why that Final Research Data likely cannot be shared under the Framework or likely cannot be shared under the Framework within the timeframe stated in the Framework. Only strong justifications will be accepted; and*
3. *include a proposal on alternative means in which that Final Research Data can be shared e.g. by anonymising or aggregating that Final Research Data, obtaining the relevant consent to disclose that Final Research Data; obtaining the relevant copyright permissions or sharing that Final Research Data at a later, more appropriate time (in which case, an extension of time should be sought pursuant to the Framework).*

*“****Final Research Data****” means, in relation to the Research, recorded factual material commonly accepted in the scientific community as necessary to document and support research findings, regardless of whether the data is used to support publications. Final Research Data that are clinical in nature shall include the human subject’s basic demographics (including but not limited to the human subject’s age, gender and race), disease and condition (i.e. the diagnosis), unless the information is not collected in the Research. Final Research Data shall 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.*  | * You are required to provide strong justifications for anticipating that you cannot share all Final Research Data under the Framework, or cannot share the Final Research Data under the Framework within the timeframe stated in the Framework. Possible justifications may include the need to protect human subjects’ confidentiality or adhere to consent agreements, or the anticipated need for more time to obtain intellectual property rights (such as patent application) over the Final Research Data and/or an invention/product that may be invented/developed using the Final Research Data or to work on a product that may be developed using the Final Research Data for public benefit.
* You shall include alternative means to produce a version of the Final Research Data that can be shared, as far as possible. (E.g. in relation to research projects with sensitive data, an aggregated version of the Final Research Data can be shared.)
* *Example A: Considering the risk of re-identification of this unique population, only aggregated Final Research Data will be available for sharing.*
* *Example B: Only anonymised patient demographics, baseline disease characteristics will be available for sharing. Post-treatment assessment data will not be available for sharing due to agreement with industry collaborator related to intellectual property rights. The team had negotiated with the industry collaborator to allow sharing of post-treatment data should consideration with intellectual property rights is no longer a concern.*
 |
| 1. **Data Use Limitation**
 |  |
| *Indicate any anticipated limitations to using the Final Research Data shared under the Framework.* *Such limitations should be minimised where possible and be substantiated with valid reasons e.g. where the limitation is due to limited consent obtained for the use and disclosure of the Final Research Data. Examples of such limitations include the use of the Final Research Data by non-profit organisations only, the use of the Final Research Data for health and/or biomedical research only, or the use of the Final Research Data for research related to a specific disease only.* | * Approved requests to access the Final Research Data will be subject to the data use limitation set out in this section.
* Anticipated data use limitations should be minimised where possible and be substantiated with valid reasons.
* *Example A: Data can only be used for non-commercial research as part of the terms of patient consent.*
* *Example B: Data can only be used for liver cancer-related research as part of the terms of patient consent.*
 |
| 1. **Institutional Review Board (IRB) Endorsement (For a Research Study that involves human subjects)**
 |  |
| *For research study that involves human subjects, please state whether an IRB has reviewed and endorsed the Data Sharing Plan.* | * Where the research involves human subjects, the Lead Principal Investigator and his/her institution shall ensure that the Data Sharing Plan is reviewed and endorsed by the relevant IRB before any work requiring IRB’s approval may commence, to ensure adequate protection of the human subjects. For the avoidance of doubt, if the review and endorsement has not been completed/given at the time that the Data Sharing Plan is submitted, you should indicate “No”.
 |
| 1. **Lead Principal Investigator’s Undertaking**
 |  |
| *The Lead Principal Investigator shall ensure that the Data Sharing Plan is (i) in accordance with the Framework, and the Lead Principal Investigator’s institution’s data sharing/management policy; and (ii) in compliance with all applicable legislation prescribing requirements in relation to data.* | * This undertaking is mandatory. The Data Sharing Plan must be in accordance with the Framework, and the Lead Principal Investigator’s institution’s data sharing/management policy, and in compliance with all applicable legislation that prescribes requirements in relation to data.
 |

|  |  |
| --- | --- |
| **Supplementary Data Management Questions** | **Examples (For Reference)** |
| 1. **What data will you collect, create or reuse?**
 |  |
| *Describe format of data (e.g., text, numeric, audio-visual, models, computer code, discipline-specific, instrument-specific) and give an indication of the anticipated volume of data. Outline and explain your choice of format and consider the implications of data format and data volumes in terms of storage, backup and access.**Indicate usage of any secondary data. This could include any existing data, data from earlier projects or third-party sources such as administrative data available in TRUST (you may refer to the TRUST webpage for more information on the data available in TRUST –* [*https://trustplatform.sg/collaborate/access-trust-data/*](https://trustplatform.sg/collaborate/access-trust-data/)*).* | * *Example A: Data will be collected using several file formats, including CSV excel spreadsheet, MS Word, and equipment specific software such as FlowJo for flow cytometry data. Graph Pad prism will be used for graphs and statistical analyses. The anticipated volume of data will be <50GB, and this should not pose any issue for our current storage and backup plans. Existing data from earlier projects will also be used.*
* *Example B: Audio recordings of interviews will be captured as MP3 files, and transcription will be stored as ASCII plain text. Digital images will be stored in TIFF or JPEG format. No proprietary software or file types are expected to be used. The anticipated volume of data is likely >1TB considering the expected number of audio and image files. Storage and backup will be planned accordingly to accommodate the volume of data. Secondary data will include data extracted from the electronic health records.*
 |
| 1. **How will the data be collected or created?**
 |  |
| *Describe data collection method (e.g., experimental (generated by lab equipment), computational/simulation (generated from computation models), observational (recordings of specific phenomena at a specific time or location), derived (produced via processing or combining other data), reference (extracted from published and/or curated datasets)).**Specify any community agreed or other formal data standards used (with URL references) to enable interoperability.* *Explain how the consistency and quality of data collection will be controlled and documented. This may include processes such as calibration, repeat samples or measurements, standardised data capture or recording, data entry validation, peer review of data or representation with controlled vocabularies.* | * *Example A: Data will be collected through experiments with protocols outlined in the proposal. Data quality will be controlled through repeat measurements, cross checking with existing measurements and using standardised protocols. Equipment will also be calibrated at regular intervals in accordance with good laboratory practice. Community agreed data standards will be used as far as possible. For example, the Minimum Information About a Microarray Experiment (*[*MIAME*](https://pubmed.ncbi.nlm.nih.gov/11726920/)*) data standard will be used for microarray data, and the Minimum Information about a high-throughput SEQuencing Experiment (*[*MINSEQE*](https://www.fged.org/projects/minseqe/)*) will be used for sequencing data.*
* *Example B: Interviews and focus group discussion will be conducted as proposed in the research proposal. Interviews and discussion sessions will be recorded using digital recorders and will be transcribed subsequently using a transcription software. A data collection protocol will be developed, and this will include questions to be raised in interviews and focus groups, as well as guidelines on recording. Data collected will be reviewed by the PI to ensure consistency. The Observational Medical Outcomes Partnership (*[*OMOP*](https://www.ohdsi.org/data-standardization/)*) Common Data Model will be used for observational data.*
 |
| **A3. How will collected data by linked?**  |  |
| *Describe how utility of collected data can be further leveraged with longitudinal data (past and present) from administrative data available in TRUST or other sources to support the project, if applicable. Applicants are encouraged to plan ahead for secondary use of datasets and data linkages (e.g., obtaining necessary consent to link data with administrative data available in TRUST or other sources and secondary use of datasets) to expand the utility of the data collected.* | * *Example A: Data collected from the project will be linked with the human subjects’ clinical data available in the National Electronic Health Record, so as to save time and cost associated with collection of the clinical data. PIs will ensure that the consent for data collection and linkage is obtained from the human subjects.*
* *Example B: Cohort data generated from this study will be linked with the GUSTO cohort datasets and the administrative data from the Health Promotion Board to create a richer dataset for comprehensive analysis. The necessary consent for data collection and linkage, and the relevant data sharing agreements will be obtained for the study.*
 |
| **A4. How will the data be stored and backed up during the research?** |  |
| *Describe where and how you will store the data during the research.**Describe the backup and archiving regime you will use to back up all your data to prevent its loss.* | * *Example A: Data will be stored in storage drive, which will be backed up in another set of storage drives bi-weekly. The storage drives will be encrypted for data protection purposes.*
* *Example B: The data will be saved to the institutional server. The server will be maintained and managed according to the institutional IT & security policy, including backup and archival.*
 |

1. The NMRC may be the sole party providing funding for the Research, or it may be co-funding the Research with other public/non-public organisations. [↑](#footnote-ref-1)