PROCEDURE

Metro South Health Research Biorepositories – Databases, Tracking, Records and Documentation

PR2017/109
Version No. 3.0

PURPOSE

Metro South Health (MSH) is committed to compliance with national guidelines and laws safeguarding the privacy and confidentiality of patients/patient/participants that have provided personal and clinical data and samples to MSH research biorepositories for research purposes. The generation of clear, accurate, comprehensive and retrievable records and documents are vital to the research biorepository's compliance and success. The purpose of this procedure is to outline general principles to ensure that the privacy of the patient/participant is safeguarded and ensure that databases, records and documents are maintained with common essential standards.

OUTCOME

Whilst research biorepositories must be operated in accordance with the MSH Research Biorepository Governance Framework, principles may be adapted so that they are appropriate to the mission and goals of each research biorepository.

This procedure applies to all MSH or Queensland Health (QH) employees whose usual reporting line is through a MSH facility or service (including visiting medical officers, visiting health professionals, students and researchers) who operate or access, or who propose to establish or access, a research biorepository that includes biospecimens collected, processed or stored within MSH facilities.

Failure to comply with this procedure may amount to research misconduct on the part of the responsible individual. This procedure must be read in conjunction with other MSH Research Management and Research Biorepository procedures.

KEY PRINCIPLES

The following key principles guide MSH research biorepositories in their databases, tracking, records and documentation arrangements. The way in which individual MSH research biorepositories put these principles into operation may be scaled in relation to the research biorepository's size of operations.

- Appropriate annotation of biospecimens is crucial to the overall usefulness of research biorepositories for scientific research.
- Research biorepositories store collected biospecimens and samples using multiple methodologies and processes. Researchers rely on banked biospecimens for a wide variety of purposes, using different platforms and technologies. The data recorded by investigators and research biorepositories depend on the types of biospecimens collected and the research projects' objectives.
- MSH research biorepositories must give consideration to privacy and confidentiality matters and issues that arise in the conduct of biospecimen collection and research. Confidentiality of personal information as well as data associated with tissue and biological samples is critical. The issues concern storage, transmission, retention and sharing of patient/participant information in a manner

compliant with legislative and ethical requirements. All personal information must be encoded as early as possible after collection however research biorepository personnel must be able to re-identify samples if/when required.

- As Custodians of biospecimens and associated information, MSH research biorepositories have a responsibility to maintain complete and auditable records. All records and documents that have to be generated and maintained as part of the operation of a MSH research biorepository including; written notebooks, original paper records, true copies such as photocopies, microfiche or microfilm as well as electronic records and documents (eg CD, DVD, USB etc) and presentations or publications of approved research arising from dispatched samples must be managed in accordance with this procedure.
- All MSH research biorepositories and to personnel involved in all aspects of tissue banking, tumour banking, biobanking and collection of biospecimens that have access to patient information, samples and research results and are involved in generating, maintaining and managing records and documents must implement systems and processes to ensure compliance with the MSH Research Biorepository Governance Framework.
- Research biorepository data records must be monitored to ensure completeness and accuracy. Custodians of biospecimens are responsible for keeping proper records of all uses (including samples that have been dispatched to researches for approved Human Research Ethics Committee (HREC) research projects or other uses by research biorepository personnel (ie recorded as 'used' due to: damaged; degradation; discard or entered in error etc) that have been made of the material, whether by themselves or by others.
- Custodians of biospecimens must ensure that all use by independent parties have appropriate MSH HREC approval, and keep copies of such approvals for easy reference.
- When encoded samples are provided to a third party, the Custodian is responsible for safe keeping of the code enabling samples to be linked back to individual donors.
- All of the resources held by the research biorepository must be maintained and tracked through an information management system that includes administrative data, the biospecimens and data derived from their analysis, phenotypic data, and any other information collected from or about the patient/participant or their biospecimens.
- Research biorepositories must meet all relevant legislative and regulatory requirements including but not limited to; *Information Privacy Act 2009 (Qld)* and *Privacy Act 1988 (Cth).*

LEGISLATION OR OTHER AUTHORITY

Legislation

- *Hospital and Health Boards Act 2011 (Qld)*
- *Information Privacy Act 2009 (Qld)*
- *Human Rights Act 2019 (Qld)*
- *Public Health Act 2005 (Qld)*
- *Privacy Act 1988 (Cth)*
- *Therapeutic Goods Act 1989 (Cth)*
- *Transplantation and Anatomy Act 1979 (Qld)*

To the extent an act or decision under this document may engage human rights under the *Human Rights Act 2019*, regard will be had to that Act in undertaking the act or making the decision. For further information on the *Human Rights Act 2019* see: https://www.qhrc.qld.gov.au/

Regulation

- Transplantation and Anatomy Regulation 2004 (Qld)

Statements, papers and guidelines

- Canadian Tissue Repository Network: Policies and Standard Operating Procedures
- Government of Western Australia: Guidelines for Human biobanks, genetic research databases and associated data
- Medical Research Council: Use of Human Samples in Medical Research
- National Cancer Institute: Best Practices for Biospecimen Resources
- National Health and Medical Research Council (NHMRC):
    - National Statement on Ethical Conduct in Human Research 2007
    - Australian Code for the Responsible Conduct of Research 2018
    - Biobanks Information Paper 2010
- Organisation for Economic Co-operation and Development (OECD)
    - Best Practice Guidelines for Biological Resource Centres
    - Guidelines on Human Biobanks and Genetic Research Databases
- Queensland Biotechnology Code of Ethics
- World Health Organisation (WHO): Common Minimum Technical Standards and Protocols for Biological Resource Centres Dedicated to Cancer Research
- **MSH policies, procedures, manuals and frameworks**
- MSH Research Management Policy (PL2017/55)

RESPONSIBILITIES

Executive Management

Must ensure all research biorepositories established in MSH are consistently operated in accordance with collaborative, harmonised, clear and detailed publicly available policies, procedures and Standard Operating Procedures (SOPs).

Metro South Research

Support Custodians in the operational arrangements of each research biorepository through the provision of guidance and support when interpreting principles and provisions contained within the MSH Research Biorepository Governance Framework.

MSH Research Biorepository Strategic Oversight Committee

Review and approve all Research Protocols and SOPs which provide clarification around databases, tracking, records and documentation for MSH research biorepositories.

MSH HREC

Ethically review MSH research biorepository Human Research Ethics Applications (HREAs) and associated documents (eg Research Protocol, PICF and Curriculum Vitaes).

Custodian/Principal Investigator – responsible officer

Ensure the research biorepository's processes for databases, tracking, records and documentation not only supports individual and institutional interests however also ensures that high quality annotated biospecimens will be available for future research efforts.

Research biorepository manager

Write, revise and update organisational and administrative SOPs pertaining to databases, tracking, records and documentation for MSH research biorepositories.

Laboratory technician/technologist assistant/clinical personnel

Research biorepository personnel must possess sufficient educational background, experience and training to assure that assigned tasks pertaining to the collection of biospecimens from MSH patients/participants are performed in accordance with the MSH Research Biorepository Governance Framework and applicable SOPs.

SUPPORTING DOCUMENTS

Attachment 1 - Application

Attachment 2 - Selection of Research Biorepository Informatics Management Systems

Attachment 3 - Validation and Operation of Research Biorepository Informatics Systems

Attachment 4 - Confidentiality Agreement for Employees

Attachment 5 - Databases, Tracking, Records and Documentation Checklist

DEFINITIONS

See the MSH Research Biorepositories Glossary

PROCEDURE - DATABASES, TRACKING, RECORDS AND DOCUMENTATION

STEP 1: Privacy and confidentiality

The Custodian must ensure privacy and confidentiality of patients/participants is protected through a combination of appropriate mechanisms.

STEP 2: Protection of data

The Custodian must ensure that the data contained within the research biorepository databases are protected in accordance with Australian legislation.

STEP 3: Identification of data

Data protection must where appropriate involve the separation of information that can readily identify an individual from other data (eg genotypic data). Informatics systems should track clinical data associated with a biospecimen and/or link biospecimen data with external sources of clinical data, where applicable.

STEP 4: Restrictions on access

The Custodian must ensure only a restricted number of authorised personnel have access to information identifying or potentially identifying patients/participants and that this is monitored and documented.

STEP 5: Infrastructure

The Custodian must ensure a robust infrastructure is in place consisting of both hardware and software components, to prevent unauthorised access. Additionally, the research biorepository must manage and store data and produce electronic catalogues based on authenticated information. The authentication of data may differ from each research biorepository however generally a research biorepository should; provide traceability of data through a history of modifications (dates and signatures of inputs, validations, modifications and deletions) and given signature for data entry, validation, modification or deletion.

STEP 7: Biospecimen tracking

The research biorepository must ensure all biospecimens are tracked through an appropriate inventory tracking system. The research biorepository must use standard terminology and formats for data management and exchange and standard processes for data transmission to networks. Please see Attachment 2 - Selection of Research Biorepository Informatics Management Systems for assistance regarding this process and Attachment 3 - Validation and Operation of Research Biorepository Informatics Systems for more information.

STEP 8: SOPs

The Custodian must ensure there are SOPs in place regarding databases, tracking, records and documentation. The Custodian must have a clearly articulated SOP of whether data will be accessed from health or other records, and/or be independently assembled, and whether or not these data will be linked with or stored in the research biorepository. Additionally, research biorepositories releasing biospecimens and/or data should have a clearly articulated SOP on whether and how the results of research and analyses carried out using its resources should be returned to the research biorepository, incorporated into its databases and how access to such results for further research will be managed (if required).

The Custodian must include SOPs for managing 'incidental findings' from researcher results/feedback and/or provide an ethically defensible plan if incidental findings are not intended to be provided to patients/participants. Please see MSH Research Management - Biospecimen Ethics and Participant Information and Consent Form Procedure (PR2017/115) for more information.

STEP 9: Confidentiality agreement for employees

MSH research biorepository Custodians must request for all research biorepository personnel (that will have access to patient/participant or research information) to complete a Confidentiality Disclosure Agreement.

Attachment 4 includes template Confidentiality Disclosure Agreement that may be used by all MSH research biorepositories.

STEP 10: Self-audit, review and compliance

Utilise Attachment 5 - Databases, Tracking, Records and Documentation Checklist to aid in self-auditing, review and compliance.

PROCEDURE DETAILS

Procedure Number
PR2017/109

Procedure Name
MSH Research Biorepositories – Databases, Tracking, Records and Documentation Procedure

Policy Reference
PL2017/53
MSH Research Biorepositories Policy

Supersedes
Version 2.0

Procedure Author
Erica Wright, Manager, Research Development, Metro South Research, Metro South Health

Portfolio Executive Director
Professor John Upham, Chair, Metro South Research, Metro South Health

Approving Officer
Professor John Upham, Chair, Metro South Research, Metro South Health

Approving Date
05 July 2021

Effective From
05 July 2021

Date of Last Review
05 July 2021

Date of Next Review05_ July 2024 (within the next 3 years)

ATTACHMENT 1 - Application

1.0 Privacy and confidentiality

The use of MSH research biorepositories and accompanying data is critical for medical research. The public and patients/participants must have confidence that research biorepositories and researchers will use and handle such material with confidentiality. Research biorepositories must not disclose to any person any information or documents whereby the identity of the patient/participant may become publicly known.

It is important to ensure that sensitive information is used ethically and optimally for the research to benefit health and knowledge. Safeguarding the privacy of the patients/ participants should be of primary importance. Rules protecting the privacy of personal information collected for research purposes are outlined in the National Statement on Ethical Conduct in Human Research 2007 ('National Statement').

Of utmost importance is whether biospecimens are practically identifiable. However, highlighting the potential identifiability of all biospecimens also highlights the need to ensure that the privacy and confidentiality of any stored biospecimens are protected. If it is assumed that it is impossible to make biospecimens completely anonymous, appropriate attention can be focused on data security and safeguards to minimise the risk of individuals being identified

1.1 Consent for the collection use and disclosure of personal information

To comply with MSH key principles on privacy and confidentiality, patients/participants must be informed about how information about them will be used. MSH research biorepositories must have each patient/participant's explicit consent to obtain, store and use information about them. Consent is required for the collection of personal information and the subsequent use and disclosure of this information. MSH research biorepositories must seek consent for the use or disclosure of the information at the time of biospecimen collection.

In keeping with the concept of 'informed consent', the research biorepository must make an effort to ensure that patients/participants are advised of the overall purposes for which their information will be used. Patients/participants must be confident that MSH research biorepositories will follow the guidance of the MSH HREC for reviewing and approving access to their material.

Information should not be used for purposes that have not been specifically identified in the consent process without seeking the guidance of the HREC of record. Please see MSH Research Management - Biospecimen Ethics and Participant Information and Consent Form Procedure (PR2017/115) for more information.

1.2 Openness about personal information processes and SOPs

MSH research biorepositories must be open about their processes and SOPs with respect to management of personal information. Patients/participants should be able to acquire information about processes and SOPs without unreasonable effort and this information should be made available in a form that is generally understandable.

MSH research biorepositories should make information about processes and SOPs available in a variety of ways. This may include brochures available at its place of business or at promotional events and online access to SOPs, forms and selected educational material.

1.3 Protection of privacy and data

The Custodian must ensure privacy and confidentiality is protected through a combination of mechanisms, as appropriate, including for example; secure storage of samples and data, coding and encryption, data enclaves, and honest broker systems.

The Custodian must ensure that the data contained within the research biorepository databases are protected in accordance with Australian law. The Custodian must ensure consideration is given to the extent to which the genetic data held by the research biorepository might allow, alone or in combination with other available samples and data, the patient/participant to be identified. The Custodian must ensure a plan is developed to manage and minimise all risks identified.

Data protection should, where appropriate, involve the separation of information that can readily identify an individual from other data, including genotypic data. The research biorepository must protect privacy and confidentiality through a combination of mechanisms including, for example: secure storage of biospecimens and data, coding and encryption of these, logging of any access to biospecimens or data, data enclaves, and honest broker systems. Where feasible, patient/participant identifying data should be encrypted from the point of collection through all phases of data handling including storage, manipulation and transfer of data.

The Custodian must ensure there are SOPs in place on protection including whether certain data will not be available for access in order to prevent the possible identification of patients/participants. The research biorepository must have in place a robust infrastructure, including equipment and software, so as to prevent unauthorised access to its databases.

The Custodian must also ensure that only a restricted number of properly authorised staff, and in accordance with obligations of confidentiality, have access to information identifying or potentially identifying patients/ participants. Such access should be monitored and documented and only be exercised when necessary.

1.4 Ensuring safeguards for personal information

The research biorepository must be established, managed, governed, and operated in such a way as to prevent inappropriate or unauthorised access to or use of patients/participants' biospecimens and personal data and/or information. Patients/participants' must be made aware of what safeguards are in place to protect their confidentiality. The security safeguards should protect personal information against loss or theft as well as unauthorised access, disclosure, copying, use or modification. MSH research biorepositories must protect personal information regardless of the format in which it is stored.

Security safeguards appropriate to the sensitivity of the personal and clinical information should protect this information. Methods of ensuring security of associated information should include the following methods:

- Physical measures such as locking research biorepository filing cabinets, freezers and restricting access to offices and laboratories.
- Organisational measures, such as limiting access on a 'need-to-know' basis.
- Technological measures, such as using passwords, firewalls, and encryption.
- Encoding procedures such as de-identification or de-personalisation of source data.
- Routine back-up of data and information stored electronically.

The Custodian must establish and implement specified SOPs for the protection of biospecimens and data, especially those potentially permitting, whether directly or indirectly, the identification of the patient/participant.

1.5 Identifying purposes for the collection of personal information

Personal and linked medical information relating to the patients/ participants and biospecimens should always be treated as confidential. The patient/participant should be made aware of the type of personal and medical information that will be used by researchers. Each biospecimen should have a unique ID assigned to it in the system. The informatics system should have the capability of linking the labels on the physical biospecimen container (eg paper labels or barcodes) to other information regarding that biospecimen in the system.

There is growing recognition that the concepts of identification and anonymity in relation to health data and in particular to research biorepositories are not easily defined and terms such as de-identified and anonymised are misnomers. The concept of personal or identifying information is perhaps best viewed as a spectrum or continuum from unidentified to identified.

The Custodian must consider the extent to which the genetic data held by it might allow the identification of patients/ participants, either alone or in combination with other available data and reference samples. The research biorepository should establish a clearly articulated SOP of whether certain data or combinations of data will not be made available and for which reasons.

An individual will be identifiable if the:

- information is identified with the individuals name, image, date of birth, address or other personal identifier
- information contains a unique personal identifier and the holder of the information also has a master list linking the identifiers to individuals
- number of different pieces of information known about a particular individual enables someone to link the known pieces and complete the (re) identification of some or all those in the data list
- person holding the information can merge or link it to other information which will enable them to identify the individual/s
- identity of the individual can be established from aggregated data because of the small number of persons within a particular category
- information derived from a sample can be used to identify an individual, or enable (re) identification.

In every case a judgement must be made as to whether the identity of an individual can reasonably be ascertained by the holder of the information. This decision on identifiability of data depends on the probability that a specific individual can be identified from the information. Equally important is that identifiability is dependent on both the amount of information held and on the skills and technology employed by the holder.

Biospecimen collections which are for specific research projects must ensure patient/participant privacy and confidentiality is protected outside of the research project team.

Separation between non-identifiable and identifiable data is technology and information based but is also highly dependent upon the ethical conduct, adherence to good governance practices and an understanding of the duties owed, by Custodians and responsible officers, in relation to the research biorepository’s associated data and other databases.

1.6 Accountability for personal information

Accountability of MSH research biorepositories for compliance to applicable legislation rests with the Custodian, research biorepository manager and/or designated official for day-to-day collection and processing of personal information. The name of the person, accountable for overseeing compliance to these principles, should be a matter of public record.

As the custodian of personal information in its possession, including information that may be transferred to a third party, the research biorepository must use contractual means (such as a Material Transfer Agreement) to ensure a comparable level of protection while the information is being used by the third party. Please see Material Transfer Agreements, Packaging and Shipping Procedure (PR2017/107) for more information.

1.7 Limiting collection

MSH research biorepositories will not collect personal information indiscriminately. Both the amount and the type of information will be limited to that which is necessary for the purposes identified by the collecting research biorepository in the consent process. Key principles which pertain to limiting use, disclosure, and retention of personal information are listed below:

- Personal information should not be used or disclosed for purposes other than those for which it was collected.
- The research biorepository should control the release of information to researchers by evaluating each request for scientific merit and compliance with approved ethical standards.
- Researchers using the research biorepository can only use biospecimens or disclose information in accordance with the terms and conditions outlined in a Material Transfer Agreement (MTA).
- A research biorepository must implement protocols (approved by the MSH HREC) with respect to the retention of personal information:
    - For cases of withheld consent, all case related tissue and data held (electronically or on paper) by the research biorepository should be removed or destroyed (if a waiver of consent hasn't been sought).
    - For cases of revoked consent, all case related unused tissue and data should be limited or destroyed.

Guidance from the MSH HREC should be used in the management of case related tissue and data accrued, that cannot be destroyed as it may already be engaged within a Research Protocol. In some cases, such material may be used as anonymous donor/tissue without information about clinical characteristics and with MSH HREC approval. Please see MSH Research Management - Biospecimen Ethics and Participant Information and Consent Form Procedure (PR2017/115) for more information.

1.8 Accuracy of personal information

To minimise the possibility that inappropriate or insufficient information may be used to make decisions or conclusions about the research undertaken, personal information and data should be accurate, complete and up-to-date.

1.9 Individual access to own personal information – incidental findings

Personal information includes data that has been collected (including lifestyle and clinical data) but not data created by research. Exceptions to individual access should be controlled by the MSH HREC and may be warranted if the information contains references to other individuals, is prohibitively costly to provide, is not traceable or cannot be disclosed for legal, security or commercial proprietary reasons. If valuable medical information becomes available from research on biorepository samples, the decision to contact the patients/participants or their families to offer benefits of that research should be guided by the MSH HREC of record and best clinical practice. If the research is likely to produce information relevant to the health and wellbeing of the person from whom the tissue was derived, procedures to allow patients/ participants to be identified for appropriate follow-up should, wherever possible, be included in the research proposal.

Where research may discover or generate information of potential importance to the future health of

patients/participants, or their blood relatives, researchers must prepare and follow an ethically defensible plan to disclose or withhold that information. This plan must take into account the clinical relevance of the research information, the types of genetic test used in the research, and the results of those tests. In addition, the plan should:

- enable patients/ participants to decide whether they wish to receive the information and who else may be given the information
- set out a process for finding out whether those other people want to receive information
- include procedures to inform patients/participants that the information would remain potentially identifiable
- include measures to protect the degree of confidentiality that patients/participants wish to maintain.

When patients/participants or their relatives are to be given or notified of genetic information that may be important for their health, the plan should either provide access to genetic and clinical advice and counselling, or clearly recommend to patients/participants that they seek these services. Such advice and counselling should be provided by professionals with appropriate training, qualifications and experience.

Where patients/participants or relatives prefer not to receive genetic information that is important for their health, they should be advised that they will be approached to confirm this decision when the results of the research are available.

Where the potential relevance of genetic information to patients/participants' health is not clear until after interim analysis of the research information, patients/participants should again be given:

- the option of being notified of the existence of that information
- the option of receiving the information
- access to, or a recommendation to seek, advice or counselling about the implications of these decisions.

Advice about the results of genetic research needs to include a clear explanation of the difference between research and clinical testing and to clarify any need for clinical testing of research results.

Identifiers of genetic material or related information:

- should not be removed without the consent of patients/participants, if removal would make it difficult to communicate personal results
- should be removed if patients/participants request it, provided they have been informed that the material or information would remain potentially identifiable.

If individual results are to be released to patients/participants, it is important to consider whether a genetic counsellor is required to assist with the disclosure or, at the very least, to be available to explain the significance of the results.

Please see MSH Research Management - Biospecimen Ethics and Participant Information and Consent Form Procedure (PR2017/115) for more information.

1.10 Release of research results to research patients/participants

In research involving humans, it is normally expected that research patients/ participants will be provided with information in general terms about the research results (assuming that the patients/participants are contactable), and that the results will be published in order to contribute to the advancement of public knowledge.

Research biorepositories have the potential to reveal medically relevant information about the health or future health of patients/participants and possibly their relatives and group or community; but the question of release of more specific research results back to patients/participants raises some difficult issues, including the management of patients/participants' legitimate expectations.

The range of potential issues involved in providing such feedback can be anticipated and addressed to some extent in the consent process; however, there are limits to the extent to which advance consent can address all the ethical and legal dilemmas that may arise in relation to disclosure of results.

From the outset, it is important to clarify the range of possible results or information that may be at issue, including results of individual patients/participants or aggregated patients/participants' results either from initial analyses or from the verified or otherwise tangible results of research using the research biorepository, as well as incidental findings. Research biorepositories need to have a clearly articulated SOP on the nature of the feedback to patients/participants and, if feedback of either interim and/or final results is planned, whether it will be in the form of individual or aggregate results. The SOP should also address how incidental findings from the research are to be handled.

There are, nevertheless, a number of reasons why release of individual results to patients/participants may be problematic and providing patients/participants with individual-level results should give special consideration to the numerous complexities that doing so presents. The results arising from research biorepository research are unlikely to be scientifically validated, or meaningful to patients/participants, and could potentially be harmful in the absence of informed interpretation, counselling and support. For this reason, it is generally assumed that release of individual results to research biorepository patients/participants would not normally be appropriate, particularly in circumstances where there is no therapeutically useful course of action available.

1.11 Complaints

MSH research biorepositories must put SOPs into place to receive and respond to complaints or inquiries about its processes and practices relating to the handling of personal information. Please see Governance, Oversight and Management Procedure (PR2017/98) and Strategic Oversight Committee and Compliance Procedure (PR2017/99) for more information.

1.12 Independent intermediaries

Best practice in genetic research involving genetic databases requires the appointment of an independent intermediary between the researcher and the data and samples (a 'gene trustee') to protect the privacy of samples and information. The value of this approach lies in the separation of any identifying information from all sensitive data and materials held in a database. No matter who obtains access to this material, they will be unable to identify it without contacting the gene trustee, who will be bound not to release any identifying information without the consent of the individual. Due to administrative costs and practicalities, MSH does not require the use of independent intermediaries however they may be utilised if/when required.

2.0 Records management

The importance of adequate data management cannot be overstated. There is an unmet need for the development of open-source software for research biorepository management. Research biorepositories must develop a record management system that permits detailed records to be made concurrently with the performance of each step in the collection, processing and distribution of biospecimens.

This may include but is not limited to: informed individual consent, procurement, processing, preservation, quarantining, testing, record review, releasing, labelling, storage, distribution and quality control of biospecimens.

The use of biospecimens and accompanying data is critical for medical research. Clear, accurate and complete records are essential to any research program. As Custodians of samples of biospecimens, research biorepositories are responsible for keeping proper records. The following principles guide MSH research biorepositories in maintaining compliant records and documents. The research biorepository must have a clearly articulated SOP on the duration of storage of biospecimens and data.

Records should be created and maintained in a manner that allows full traceability. Laboratory Information Management System (LIMS) allows the management of such data. Data security systems must be adequate to ensure confidentiality and safety. Record management must be regularly audited and records must be kept for at least ten years after expiration of biospecimen storage or biospecimen distribution. Electronic records must be adequately protected (regular backups on appropriate media, intrusion-proof management systems).

2.1 Records and documentation

Recent developments in molecular biology and genomics have essentially enhanced the value of clinically annotated human tumour tissue in translational research and drug discovery. Adherence to best practices in the generation and maintenance of complete and accurate documentation is important in ensuring the value and utility of resources within the research biorepository.

Intellectual property rights can only be protected adequately if all records and documents are thorough, accurate and contemporaneous.

Research biorepository documentation for scientific networking is also an important aspect. Each MSH research biorepository must be registered on the Metro South Research Internet Site to inform the scientific community on the nature of the collection and of its general content.

Research biorepositories should track researchers' requests for biospecimens with specific clinical data to guide the refinement of clinical data collection, as appropriate, based on the intended purpose of the resource and if the research biorepository is the point of access for biospecimens and associated clinical data. Research biorepositories should routinely summarise this information and provide it to an entity that maintains and/or collects the clinical data in order to improve the collection of clinical data.

2.2 Notebooks and electronic records

All raw data or personal and clinical information may be recorded and retained in laboratory notebooks or in an electronic database dedicated to that purpose. Machine print-outs, consent forms, questionnaires, chart recording, autoradiographs, forms, letters etc which cannot be attached to the main record should be retained in a separate ring-binder/folder that is cross-indexed with the main record.

Notebook and electronic records should be entered as soon as possible after the data is collected or generated. Recorded data should be identified by date of the record and date of collection if the two do not coincide. Subsequent modifications or additions to records should be clearly identified and dated.

Research biorepositories must implement processes for quality assurance of data collected and recorded electronically.

Where feasible, internal annotated digitised data/images should be recorded and retained in a "raw" or original format as well. This is especially relevant where data/images undergoing digitisation are subsequently enhanced. If possible, both the original and enhanced forms should be stored. All electronic records should be backed-up regularly; duplicate copies should be held on disc in a secure but readily accessible archive.

2.3 Personnel and users access to information and records

Access to data should be given to users on a 'need to know' basis. Users should be granted access to specific data records that they need in order to perform their duties. This access should be removed when the activity is completed.

2.4 Transmission of information and data

Information from incoming sources (such as other South East Queensland research biorepositories and/or regional Queensland collection sites) must be transmitted in a secure manner (ie flat file formats) (sent by email, password protection (sent on a different day) and/or other platforms which require secured sockets (data encryption). Outgoing data must also be transmitted in a secure manner (ie data enclaves or honest broker systems).

2.5 Physical storage of information and data

Data and records must be stored securely and with appropriate contingency plans. An example of this may include daily backup of all electronic database information (may be set up automatically in servers/backup systems). In cases of corruption, data loss is limited to 24-hour window which allows retrieval back to particular dates and times. Data and records should also be stored in a manner to permit retrospective audit if needed. Additionally, records and backup discs must be stored to maximise protection from factors such as flooding, fire or theft.

2.6 Retaining information and data

Retention of accurately recorded and retrievable information, data and results is essential for the running of a research biorepository. Information, data and results must be retained for a minimum of ten years however MSH recommends records be retained indefinitely (due to the length of some stored samples) or as long as the research biorepository Custodian deems them to be pertinent.

Custodians must decide which data and materials should be retained, although in some cases this is determined by law, funding agency, publisher or by convention in the discipline. The central aim is that sufficient materials and data are retained to justify the outcomes of the research and to defend them if they are challenged. The potential value of the material for further research should also be considered, particularly where the research would be difficult or impossible to repeat. Prior to destruction of any information or data the MSH HREC must be consulted. Researchers (who are leaving an establishment) that generated data and who wish to retain anonymised data/copies of data for future use must get specific permission to do so from both the research biorepository and from the HREC. Where personal data are involved, the request should be refused unless it is clear that future use will be consistent with the terms of the consent and contract with that researcher. A Material Transfer Agreement (MTA) governs this transaction.

3.0 Databases and systems

Appropriate design elements providing for compatibility and interfaces must be incorporated when creating the databases. The research biorepository Custodian must ensure consideration is given to using standardised approaches for the collection, storage and analysis of samples as outlined in the MSH Research Biorepository Governance Framework and/or data to facilitate cross-research biorepository data exchange and sharing.

3.1 Data management and informatics security

The value of the biospecimen for research purposes is greatly enhanced by accompanying personal or clinical data related to the individual providing the sample. Personnel should treat any information about the individual, however derived, as confidential.

The protection of personal information and individual data associated with biospecimen collection is a fundamental requirement of a research biorepository. This should be achieved through the use of safe, structured bioinformatics systems. The mechanisms of access to these systems, as well as the permissions, should be clearly defined. Backups should be made on a regular basis to avoid data loss. The communication to third parties or authorities of data files containing personal information and identifiers is strictly prohibited. Personal identifiers should be coded, and all individual data archived in the research biorepository management system should be protected with the same stringency as patient clinical files.

3.2 Sample records

MSH must know which samples, which have been attained from MSH patients/participants, are held under which facility or university HREC approval and the terms of consent. For coded samples, systems must be in place allowing research biorepository personnel to determine how samples can be used to ensure that the patient/participant's wishes can be met. This means that all samples stored within the research biorepository (including those that are transferred from other places) are recorded, and their ethical approval and consent status are traceable.

Sample and data protection may be achieved via different approaches and mechanisms, and often through the combined use of various approaches. Some examples include:

- The coding and encryption of samples and data.
- Limiting access to the collection of samples and data.
- Implementation and maintenance of security measures to block unauthorised access.
- Data enclaves which involve the use of secure or controlled access databases or websites. These allow the Custodian or a third party to physically and electronically control and monitor the use of the research biorepository's database(s) by external users to ensure it complies with the terms of access and conforms to the patient/participant's consent.
- Honest broker systems which involve an independent third party who is responsible for ensuring the separation of identifying information from other data. An honest broker system may be, for example, a data protection authority).
- Where samples and data are collected by more than one research group, then each group could use their own code with none of them holding all the codes.

The Custodian must ensure SOPs are established to safeguard the privacy and confidentiality of patients/participants, samples and data, especially those that may allow, directly or indirectly, the identification of the patient/participant. Relevant SOPs must reflect the privacy risks for patients/participants that may develop as technology changes.

Additionally, the Custodian must consider; the risks that could result from any plan for the samples or data to be linked with other data sources, such as the integrated Electronic Medical Record (iEMR), the sharing of information with researchers and if this is provided de-identified and any risks of identification that may arise from research on rare conditions or conditions that are associated with sub-groups of the population.

3.3 Restricted Access

The Custodian must ensure the research biorepository is established, managed and governed in such a way as to prevent any inappropriate or unauthorised, access to or use of patients/participants' samples and data. The Custodian must ensure a robust infrastructure is in place, consisting of both hardware and software components, so as to prevent unauthorised access to databases.

Additionally, the Custodian must ensure that only a restricted number of authorised staff have access to information identifying or potentially identifying patients/participants, that such access be monitored and documented and only be exercised when necessary for carrying out research biorepository related functions. Access to the computerised inventory system must be tightly controlled. Where relevant, passwords must conform to the minimum MSH standards regarding password length, strength, life cycle, recycling etc.

In LIMS systems, security roles with defined privilege levels should be assigned to individual users of the system. Some individuals may be able to view biospecimen availability whereas others can enter or modify biospecimen descriptions and make requests to have biospecimens shipped from the research biorepository. The system may also provide a mechanism to log off users after a specified period of time during which the system is idle.

All database access attempts should be logged with the date and time of login and logout. Any failures to access the database should be logged with the date and time and reason for failure. The system should lock out a user after a specified number of failed attempts to access the system. Systems should provide vital system statistics and audit logs of all access in the database.

The inventory system must provide for single system sign-on utilising the operating system's user name and password if possible. User authentication information should always be encrypted. All remote communication should be able to be conducted on an encrypted socket (ie via a port that requires data encryption to prevent inappropriate access to secured data). For example, web-based systems should be able to implement data encryption using a Secure Socket Layer (SSL) at the browser level. All protected health information must be secured within the database through access controls and/or encryption.

3.4 Integration and interoperability

Within modern research biorepository informatics systems, integration and interoperability are highly desirable. Systems should be able to integrate with other local applications such as electronic medical records, cancer registries, pathology systems and freezer temperature monitors. This allows other systems to be the single source of truth for appropriate data.

Integration and interoperability have many benefits which include, but are not limited to, the following:

- Reduced re-entry of data. Every time data is manually re-entered from one system to another there is a risk of error. Re-entering of data can be costly.
- Data errors found and corrected in the single source of truth system should be replicated to other systems.

It is acknowledged that in some circumstances a LIMS may not be able to integrate with other systems therefore copy/paste functionality should always be employed to minimise errors eg transferring information from AUSLAB to the LIMS.

Data should be electronically convertible into formats that can easily be shared among collaborating institutions, where possible and appropriate. The informatics systems should utilise data elements from a common metadata repository. Even if the systems utilised non-standard data elements for storage internally, the system design should allow for configurable translations to one or more established standards. The inventory management system should enforce all data integrity, security and audit trail requirements for external access. To achieve interoperability, inventory management systems should do the following:

- Have a public documented application programming interface to enable other systems to integrate with it. Changes to this interface should remain backwards compatible as much as possible in order to minimise disruption for connecting systems. The application programming interface implementation should include both automated conformance and interoperability testing to ensure robustness.
- Use common public vocabularies for relevant data points (eg World Health Organisation (WHO) codes).
- The application programming interface should enforce all business and security rules on connecting systems.

Research biorepository informatics management systems should be capable of sharing appropriate, de-identified biospecimen data to users at remote locations for multiple purposes including satisfying reporting and regulatory requirements as well as searching for potential biospecimens for a proposed scientific research project. If the results data is stored in the research biorepository information management system, then it must adhere to all of the criteria listed above.

3.5 Quality control and assurance

Quality control and assurance measures must be in place to ensure, security and confidentiality during collection, storage, handling, distribution and destruction of the samples and data. In order to provide high-quality information to serve the tracking system, standards, processes and SOPs must be used to ensure and maximise the quality, objectivity, utility and integrity of the data. Periodic reviews of data quality issues and adjustments to programs and processes will ensure continuous quality improvement.

The electronic inventory system should comply with industry-applicable Current Good Practices (cGP) guidelines.

An established Quality Assurance program for the inventory system must be primarily directed at prevention of non-conformances as well as detection, corrective action and process improvement implementation. Regular Quality Assurance audits and reviews must completely document:

- User requirements, as well as industry-specific certification requirements.
- Details of the review and approval process for software developed in-house, or obtained from a third party.
- Procedures followed to test the software functionality, compared with user requirements.
- Corrective actions or processes used to handle program errors and modifications.
- Training provided to personnel associated with the use (and development, if applicable) of the inventory system.
- A periodic audit of the database should be performed to ensure accuracy of data.

Please see Quality Management System (Assurance and Control) Procedure (PR2017/110) for more information.

3.6 Custodianship of tissue data

Custodians must bear responsibility for the samples and data in their collection so that research biorepositories will be able to safeguard the interests of the patients/participants. Custodians of the tissue samples bear responsibility for keeping proper records of all uses that have been made of the materials, whether by themselves or others. If transfer of material occurs, appropriate Material Transfer Agreement processes be followed and documented. Please see Material Transfer Agreements,

Packaging and Shipping Procedure (PR2017/107) for more information. Custodians of "Existing Collections" must ensure that they make optimal use of the resource they control and seek the advice of the HREC through periodic (eg annual) review.

3.7 Backup

Regularly scheduled backup procedures are an important security function that will enable the inventory system to be restored in the event original data is lost or corrupted, most typically due to drive or other hardware failures. The database must be backed up on a regular basis, depending on the institutional policies and frequency of data modification. The more frequent the data is changed, the more frequently the backups should be made. Off-site storage is desirable. Data archives should be maintained in accordance with the maintenance of the biospecimen storage SOP. The support of these archives should be regularly updated according to its physical characteristics (obsolescence) and to software compatibility.

SOPs to preserve the integrity of IT data should include (but are not limited to) steps to limit the extent of the destructive event, protocols for periodic backing up and storing of information, procedures for off-site storage of backup data, and protocols/procedures for restoring information from backed up media. SOPs must specifically address the recoverability of information. Backups should be validated on a regular basis to ensure the data can be accurately recovered. Backup files should be stored in secure cabinets.

SOPs must specifically address the physical environment and equipment. Changes to hardware and software commonly require review and re-evaluation of these documented SOPs.

3.8 Data linkage

Where research biorepositories intend to access data from health or other records, patients/participants should be duly informed in advance, where applicable at the time of consenting, about what types of data will be extracted from such records, by which entity, through which processes, and for which purposes the data will be employed.

For access and use of such health and other records, the patient/participant's consent should be obtained, unless waiver of consent is obtained from a HREC or an appropriate authority, in accordance with applicable law and ethical principles pertaining to the protection of human subjects. Please see MSH Research Management - Biospecimen Ethics and Participant Information and Consent Form Procedure (PR2017/115) for more information.

SOPs related to data from health records should also address the issue of secondary use of health and other records, especially when combined with other data.

3.9 Terminology and formats

Where possible MSH research biorepositories should use:

- standard terminology and formats for data management and exchange
- standard processes for data transmission to networks (state, national and international networks).

Guidance may be sought from the MSH Research Biorepository Strategic Oversight Committee in relation to the implementation of the best practice principles for terminology and format specifications outlined below:

- Select data format, data representation and data transportation taking into consideration existing standards for data processing.
- Check vocabulary against standard reference lists of thesauri.
- Keep consistency among research biorepositories for searching and receiving information from catalogues and databases:
    - Each biospecimen record should contain a Minimum Data Set, a Recommended Data Set and/or Full Data Set in accordance with domain specific criteria (please see Acquisition, Attainment and Recruitment Procedure (PR2017/102) for more information).
    - Spell checking for every field should be a basic requirement.
    - International English should be choses for a preferred language of data.
    - A standardised approach should be adopted to certain scientific symbols (to avoid errors due to incurred reading of a character set, standard ASCII alternatives to symbols should be used: eg Greek letters cannot be used, they should be fully spelled (write alpha, gamma, betta…); the ^{o}symbol for temperature is to be omitted entirely (eg 37C replaces 37^{o}C); no subscripts or superscripts are allowed (eg cm3 replaces cm^{3} and CO2 replaces CO^{2}).

Research biorepositories should adopt processes to detect errors in data to improve their quality and consistency. This is an essential part of information management and should be both applied to the input of new data as well as pre-existing information in current databases:

- For existing data, a series of checks should be carried out to ascertain the validity and completeness. As more research biorepositories become associated, more searches should be made for common classes of error to allow more efficient error correction.
- For new data, wherever possible, inputting should be checked against authorised lists of not only scientific names but also thesaurus/ontology to prevent errors such as mistyping.
- Research biorepositories should present evidence that they have applied recognised protocol appropriate for each data element.

3.10 Longitudinal clinical data - data types

Research biorepositories may collect and store longitudinal data following applicable informed consent and authorisation requirements if required by the research project design and objectives.

Based on these requirements, information linked to biospecimens may include demographic data, lifestyle factors, environmental and occupational exposures, cancer history, structured pathology data, additional diagnostic studies, information on initial staging procedure, treatment data, and any other data relevant to tracking a research patient/participant's clinical outcome.

Different research biorepositories may require more or less detailed annotation based on the primary intended use of the biospecimens. The dataset for clinical annotation should be based on the needs of the research biorepository users, as well as overall feasibility, particularly for to biospecimens collected from clinical trials. Databases developed for longitudinal research projects should use coded data associated with biospecimens but should maintain a secure link to identify the research patient/participant to allow additional longitudinal data to be obtained, if permitted by law and by the research patient/participant's consent/authorisation.

SOPs must be in place to facilitate access to uniform longitudinal data (eg treatment and outcome information, as appropriate) while protecting research patient/participant's privacy and confidentiality. To collect high-quality longitudinal information, research biorepositories should ensure that dedicated and trained personnel curate longitudinal clinical data with validation of the collection process and quality assurance and control. This includes inventory functions, tracking all phases of biospecimen acquisition, processing, handling, quality assurance and control, biospecimen quality measurements (such as RNA Integrity Numbers), and distribution from the collection site (research patient/participant) to utilisation (researcher).

4.0 Biospecimen tracking

To make certain that biospecimens can be tracked accurately from the site at which they are collected through their arrival and subsequent distribution from a MSH research biorepository, effective tracking systems must be in place. Critical components of these systems include the use of unique biospecimen identifiers, appropriate biospecimen labels, electronic data inventory systems for biospecimen tracking and other features that are described in detail below.

4.1 Inventory systems

A computer-based inventory system must be in place to track the location and pertinent annotation of every biospecimen in the research biorepository. The system should also track significant events such as sample thaws, receipt and/or processing delays, destruction, processing, transfer of the sample within the research biorepository, and biospecimen distribution and return (if applicable). Full query capability for all data stored should be provided.

Biospecimens should receive a unique identifier that is linked, preferably electronically, to the uniform diagnostic data set. The use of biospecimen tracking systems for transportation, location within the laboratory, archival storage and inventory management is strongly supported. The linkages to patient data (outlined above) should be managed via de-identification/re-identification programs to protect patient privacy.

The system must have the capacity to assign a unique ID to each biospecimen entered in the database and track its lineage (parent sample to child to grandchild etc). If an ID exists, but a new biospecimen ID reflective of the inventory system into which the biospecimen is being entered needs to be added, the inventory system should be able to track the original biospecimen ID as well as the newly assigned one.

Biospecimens received in the research biorepository should be given a printed label with a barcode and biospecimen ID, if one is not already on the biospecimen container. In any recording system the usage and fate of samples must be noted (ie which projects they are used in) whether they are transferred to collaborators, disposed of (justification needed), or used up during the research. Please see Disposal, Lab/Fridge Merge and Closure Procedure (PR2017/105) for further information.

4.2 Biospecimen location

Each freezer, refrigerator or room temperature storage cabinet must have a unique identifier. A convention must be established for numbering shelves, racks, boxes, as well as each location within the storage container. Each location combination (eg freezer, rack, box, row, and column) must uniquely identify a location in the research biorepository. The inventory system must support different storage environments for the same research biorepository and also should record the container type (eg vial, straw).

When utilised by a MSH research biorepository, sophisticated inventory systems should be able to report on available storage space and able to assign and reserve space for incoming biospecimens.

To validate biospecimen location, a randomly generated biospecimen number/list or other appropriate randomised system should be checked on a subset of the samples on a regularly scheduled basis. This will ensure that the correct biospecimens are in the location specified by the inventory system (database). Certain storage systems (eg straws) are stored in containers such as goblets and the database should be able to capture such containers that do not follow the normal box/row/column configuration.

4.3 Informatics – collection and handling

The informatics system employed by a research biorepository must provide appropriate facilities for information management, linkage and exchange if the research biorepository.

The database may contain either information relating to strains held by a research biorepository (which at least should be retained as long as the strain remains viable), or other relevant data items or composite data needed by the research biorepository (eg user records).

This principle is similar to cell libraries that are made by approved researchers for specific research projects. Custodians of biospecimens collected for specific research projects should implement SOPs to control further researcher or third-party access to cell library use for other projects other than that initially approved in the HREC approval. This may be required whenever 'cryopreservation' samples (viable cells) are approved for researcher use. Additionally, a SOP must be implemented regarding the confirmation of destruction of cell libraries at either the end of the research project and/or after publication.

On the loss of a strain the database record should be either printed and stored on file or copied to a digital archive before the entry is removed from the working database, placed in reserve or annotated to indicate that it is no longer available as living material.

Annotation data (eg person, lifestyle, diagnosis, laboratory, clinical and research generated) should be accurate, quality-controlled and standardised as far as possible. Data collected may contain common data elements from the following categories including, for example:

- personal
- longitudinal clinical and diagnostic information
- treatment and outcome information
- sample information
- lifestyle and family history.

Computerised inventory and bioinformatics systems used to handle and store annotated data should:

- be responsive to the needs of multiple users
- be available for a long period of time
- use standardised terms and List of Values (LOV) to categorise biospecimens and enter data, across MSH research biorepositories (where possible and agreed upon by the MSH Research Biorepository Strategic Oversight Committee)
- use an automated data extract system or permit multiple checks of data entry to ensure accuracy
- have the ability to feed back or link standard research results and genomic and proteomic results into the system
- allow for dissemination of information to others as needed
- be searchable at varying levels for certified users

- provide security and access control to ensure privacy rights are protected
- have an inventory management system
- support integration and expansion if needed
- have maintenance features and backup capabilities.

The research biorepository must preferably choose standard data schema and protocols to make the databases distributed and interoperable. Confidential data should be clearly identified in relation with user authentication capability, encryption techniques and other related information security tools.

4.4 Biospecimen annotations and data collection

It is recommended that MSH research biorepositories adopt standardised systems for annotating the characteristics of collected biospecimens as well as data on the patients/participants or individuals who are the source of these biospecimens. The nature and extent of data collection may vary depending upon the nature and purpose of the research, the type of cancer and nature of biospecimen collected. For informatics purposes, a sample refers to a physically distinct biospecimen usually stored in a single container.

Multiple physical parts created by extraction, division into aliquots, or other physical division of a biospecimen are considered new biospecimens and are referred to in this document as samples, sometimes referred to elsewhere as derived (or child) samples, each requiring a new identifier. The origin of each sample should be recorded. Research biorepositories should define standard terms for all lineages of biospecimens, from initial collection to subsequent divisions and extractions. Where possible, biospecimen resources should employ an existing standard terminology or modify an existing standard to harmonise data elements for semantic interoperability.

4.5 Additional biospecimen descriptors

The inventory system must track biospecimen type; vial or container type; volume or size; date and time of biospecimen removal from patient/participant, collection, receipt and/or processing; processing method; storage temperature; preservatives and any other characteristics needed for the collection.

Information must be included on the history of sample processing and movement, including the location of shipments to and from external sites. Finally, any information about the sample being compromised in any way must be recorded and available to the user.

4.6 Additional information for research biorepositories

If approved by the MSH HREC, in addition to the information regarding biospecimen location, information relating to the following data sources may be maintained for a research biorepository depending upon the nature, purpose and type of the resource (if relevant, available or not stored in another interoperable information management system):

- Local coded personal identifier.
- Patient/participant information: Age of patient/participant at the time of collection, sex, race, ethnicity/language spoken or place of origin/ birth of parents/grandparents, occupation, place of residence (city/region/country) etc.
- Disease status (normal, cancerous).
- Tumour topography according to the International Classification of Disease – Oncology (ICD-O 3rd edition) and tumour morphology according to ICD-O 3rd edition.
- TNM staging and tumour grade (if applicable).

- Diagnosis: Anatomic site (eg breast), tissue type (eg normal), diagnosis (eg fibrocystic disease) and modifiers to provide additional detail regarding the diagnosis. It may be important to document the gross diagnosis (what the biospecimen was thought to be when collected, eg breast-normal), the pathological diagnosis (diagnoses rendered by pathology for the actual resection, eg breast-malignant-adenocarcinoma-ductal), quality control diagnosis of the specific sample obtained for research (eg breast-normal-fibrocystic changes) and the pathologic stage at time of surgery. In some situations, it may also be appropriate to provide the diagnosis code (ICD10) and the text of clinical diagnosis.
- Diagnostic date (if applicable)
- Diagnostic description (if applicable).
- Diagnostic procedure: Type of procedure (eg surgery), date of procedure, procedure details (eg mastectomy), procedure identification number (eg surgical pathology number).
- Type of treatment (eg chemotherapy, radiation, hormonal, immunotherapy, anti-inflammatory) prior to biospecimen collection, amounts and dates (if known).
- Information on exposure and risk factors (if applicable).
- Medication or drug history: Drug name, dose/frequency, date started.
- Family history: Relationship, diagnosis, age at diagnosis.
- Smoking history: Smoke type, smoke years, date quit.
- Vitals: Height (cm), weight (kg), alcohol history, recreational drug history, special diet, date of last menstrual period, date last follow-up, disease status at follow-up, cause of death.
- Clinical laboratory values (eg calcium, haemoglobin, etc).
- Availability of other biospecimens (eg normal vs. diseased tissue, other tissues, blood, buffy coat, serum, plasma, paraffin embedded tissue, H&E slide, formalin fixed tissue, DNA, RNA, urine, faeces, saliva, ascites fluid and synovial fluid) from the same patient/participant.
- Involvement in clinical trial/cohort study.
- If appropriate, information on medical history, treatment with past and current medication and response to therapy, concomitant disease, secondary tumours/Laboratory data.
- If appropriate, information on perioperative medication and treatment should be included as optional information:
  - protocol name
  - prior treatment
  - treatment type (surgery, chemotherapy, radiotherapy)
  - treatment start date
  - treatment end date
  - treatment response date
  - treatment response type (none, low, average, good, complete).
- If appropriate, information on duration of follow-up and disease outcome:
  - relapse date
  - relapse type (localised, distant)
  - death date
  - duration of global survival
  - duration of survival without relapse.

- Reference to the informed consent and its scope.
- Hazard status.

Whenever coded links with the patients/participants' clinical files are maintained, the annotations can take the form of "yes" or "no" answers in relation to the informed consent in the clinical file. Data may be sourced from relevant health records such as the iEMR Digital Hospital system, pathology or from the clinician.

If additional information/data sourced is not part of the research biorepositories' approved Participant Information Consent Form then further guidance must be sought from the MSH HREC. In some circumstances a 'mail out' survey may be sent to patients/participants for an approved research project when initial information was not sought during the consent process. This may include a request to access information such as current treatment, how the patients/participants' disease responded to treatment, ethnicity, place of origin, birth place, race, and languages spoken etc. Some 'mail out' surveys may also be required if additional checks a required such as making contact with other national databases (ie mortality status).

An inventory system may also be designed so that digitally scanned documents are included such as pathology reports, H&E slides of tissues collected, clinical lab reports, patient/participant consent forms and Material Transfer Agreements. The information stored will vary according to the purpose, nature, and intended uses for the biospecimen collection. Since a research biorepository may track samples of many different research projects, consideration should be given to what the inventory database can contain and what should be stored in an external database and linked to the inventory.

4.7 Additional information for non-human biospecimens

Many items listed above can also be applied to the information tracked for non-human biological samples/biospecimens, specifically animal samples, as well as additional information not included above. The following information is important when collecting for animal/fungal/plant/microorganismal biobanks or environmental biospecimen banks (ESBs):

- Collection location: (latitude and longitude [and method of determination], altitude/depth, country/area/site/city, ocean/sea/bay, indoors/outdoors.
- Collection/gathering method, in whole biospecimens; also killing method.
- Collection conditions: weather, habitat/ecosystem characteristics, association with other species.
- Taxonomic and biospecimen information: life stage, sex of the biospecimen (where applicable), species/subspecies of biospecimen and/or higher taxonomic rank.
- Taxonomic detail focusing on wild organisms: taxonomic species authority, date the identification was made and name of identifier, previous identifications, (morphological) voucher ID and collection code of the biospecimen if sample is voucher-referenced in a natural history collection, type of association to this voucher (same biospecimen / same in-situ population /etc), if biospecimen has been DNA barcoded, BOLD (Barcode of Life Database) IDs.
- Time tissue was removed from organism or site location, processed and frozen for storage.
- Processing location if divergent from collecting location, type of instrument used to collect and sub-sample, if necessary (eg stainless steel blade, titanium knife, type of container used to store the sample (eg Teflon®, polypropylene, glass, stainless steel).

4.8 Audit trail

Sophisticated inventory systems should include a full audit trail of changes made to the database; including but not limited to all biospecimen data, system metadata, and clinical data. This includes recording changes to both biospecimen data and system metadata. The audit trail must include but not be limited to: the original data; the changed data; who made the changes; how the change was made, date and time of change, and if possible, why the changes were made. This audit trail should be automatically recorded and available for read-only access.

Relevant SOPs should be developed to ensure changes to Excel Spreadsheet and databases inventory systems are recorded.

Record changes should not obscure previously recorded information in the audit trail. Such audit trail documentation should be retained for a period at least as long as that required for the electronic records and should be available for MSH review and copying.

The research biorepository inventory should be checked as part of quality assurance and quality control programmes at regular intervals (eg every two years) to assess the concordance between stored biospecimens and records. The specific position of every stored aliquot should be tracked. Each freezer, refrigerator or room temperature storage cabinet should have a unique identifier.

A convention should be established for numbering shelves, racks and boxes as well as each location within the container. The research biorepository database should be updated each time a biospecimen is moved within or out of the research biorepository.

4.9 Reporting

Sophisticated inventory systems may have the ability to produce reports to support the research biorepository workflow, document adherence to standards and practices and provide any business metrics required by the research biorepository.

The system should provide the user with an interface for specifying display content and search criteria for the report. The exact nature of this interface can vary from full “what you see is what you get” report designers to simple field selection for tabular reports. The query editor can also be presented utilising several approaches, including simple data query forms, query by example screens, customised query builders and text areas for native query specification.

The inventory system should have the ability to save reports for future execution. The inventory system should have the ability to generate report output and electronic data files (eg in ASCII, XML, or Excel format). The system should provide full access to the database for reporting, provided that the system’s security rules are enforced. This access will allow users to generate reports on inventory status, freezer status, user access, audit trail entries, and other data tracked by the database to meet their needs.

Specific research project collections may implement SOPs which detail reporting mechanisms for inventory management. If the database contains protected health information records, the security model must restrict reporting on confidential data to only authorised users. Additionally, the research biorepository should maintain SOPs about the generation, use, and destruction of reports that contain protected health information to ensure that donor confidentiality is maintained.

4.10 Validation

A closed system is defined as an environment in which system access is controlled by persons who are responsible for the content of electronic records that are in the system. Persons who use closed systems

to create, modify, maintain, or transmit electronic records must employ SOPs and controls designed to ensure the authenticity, integrity, and, when appropriate, the confidentiality of electronic records, and to ensure that the signer cannot readily repudiate the signed record as not genuine. Such SOPs and controls should include the following:

- Validation of systems to ensure accuracy, reliability, consistent intended performance, and the ability to discern invalid or altered records.
- Protection of records to enable their accurate and ready retrieval throughout the records retention period.
- Use of operational system checks to enforce permitted sequencing of steps and events, as appropriate.
- Use of authority checks to ensure that only authorised individuals can use the system, electronically sign a record, access the operation or computer system input or output device, alter a record, or perform the operation at hand.
- Use of device (eg terminal) checks to determine, as appropriate, the validity of the source of data input or operational instruction.
- Determination that persons who develop, maintain or use electronic record/electronic signature systems have the education, training and experience to perform their assigned tasks.
- The establishment of, and adherence to, written SOPs that hold individuals accountable and responsible for actions initiated under their electronic signatures, in order to deter record and signature falsification.
- Use of appropriate controls over systems documentation including:
    - Adequate controls over the distribution of, access to, and use of documentation for system operation and maintenance.
    - Revision and change control procedures to maintain an audit trail that documents time-sequenced development and modification of systems documentation.

4.11 Labels

Each biospecimen should be assigned a unique identifier or combination of identifiers, such as a number and/or barcode, which should not be reflective of its identity (ie current storage location position, clinical data, patient identifiers etc). Each biospecimen should receive a label that tightly adheres under all projected storage conditions. Information printed on labels should be resistant to all common laboratory solvents. Labels should contain an ID linking to a database containing details about the biospecimen collection and processing information. Flexibility should be allowed in the location of the label to allow for label legibility on a wide variety of containers. Please see Collection, Processing, Handling and Retrieval Procedure (PR2017/104) for more information.

Material used in composition of containers for some biospecimens may pose special problems for label adherence and therefore in some cases, the label should be able to adhere to itself. The adherence of labels to containers as well as the use of particular types of ink should be tested under conditions more extreme than the anticipated storage and processing conditions before they are put into regular use.

4.12 Labels for biospecimens

Biospecimens should be labelled in such a way that protects privacy and confidentiality and is in compliance with applicable laws and MSH policies. Biospecimens should be labelled with a unique code or ID not derived from information about the patient/participant. No other research project or personal health information should be encoded in the biospecimen ID.

For all biospecimens, the research biorepository's unique identifier for each biospecimen should be printed on the label in both barcode format and human readable form. The ID should not be reflective of its storage location in the research biorepository, as locations may change over time. Please see Collection, Processing, Handling and Retrieval Procedure (PR2017/104) for more information.

4.13 Barcoding

Whenever possible, labels should be printed with a barcode that uniquely identifies the biospecimen. Linear (1D) barcodes are adequate for small values and/or larger labels. Under some circumstances, two-dimensional (2D) barcodes are necessary. This 2D code may have to be entered as an additional code for purpose built databases that already automatically provide a unique code for individual samples and also those that provide automated storage positions (based on business rules).

Similarly, the 2D barcode tubes need additional equipment: plate readers, specific racks and boxes. Additionally, the size of the 2D tubes may cause issue for some research biorepositories depending on collection volumes. Whilst a batch of unique numbered 2D barcoded tubes can be purchased, the sequence of numbers provided is only guaranteed per order. The unique code numbering system that is used for these tubes is not able to be adapted ie the code cannot be tailor made for a lab.

2D barcodes have the advantage that scanning error rates may be lower, more information can be included on the label and they may be optimal for use on smaller vials. Some containers can be ordered pre-printed, such as straws and small vials that fit in 96-position racks. In certain situations, pre-printed containers can have an ID structure that alleviates applying another label onto the container and save supplies and labour. Each aliquot or container should be labelled with a unique barcode/number.

The MSH Research Biorepository Strategic Oversight Committee may make a determination regarding standardised barcoding/tube plate formats to be utilised across MSH. Please see Governance, Oversight and Management Procedure (PR2017/98) and Strategic Oversight Committee and Compliance Procedure (PR2017/99) for more information.

4.14 Annotations on stored biospecimens

Since a research biorepository may track samples of many different research projects or from different collections, consideration should be given to what the inventory management database can contain and what should be stored in an external database and linked to the inventory via a Unique Identification Number (UID).

In the case of biospecimens, consideration should be given to storing confidential patient/participant clinical information separately from inventory data such as sample information and location.

The informatics system may also be designed to handle digitally-scanned documents related to the sample. Relevant documents may include pathology reports, clinical lab reports, consent forms, Material Transfer Agreements or necessary permit documentation.

The following criteria may be recorded for better characterisation of biospecimens:

- Storage centre identification.
- Local research biorepository inventory code.
- Nature of biospecimen.
- For solid tissue, tissue condition

(tumour/non-tumour/interface).
- Packaging.
- Number of aliquots/quantity of biological material.
- Preservation protocol.
- Time elapsed between tissue removal and preservation (if applicable).
- Date of biospecimen collection/storage.
- Record of storage incidents.
- Temperature during transport.
- Storage temperature.
- Storage conditions (agents added to the sample).
- Documentation on processing method.
- History of freezing/thawing.
- Amount of tissue collected, and amount left over in storage.

4.15 Shipping log

Each research biorepository must maintain a shipment log to record the receipt and dissemination of shipments sent from the research biorepository (if required). This log should be integrated into the functionality of the inventory management system described above. Each shipment entry should be given a unique shipment ID. The electronic log should be able to track the following elements:

- Shipment/Invoice ID.
- Source.
- Destination.
- Date shipped and date received.
- Courier name.
- Package Tracking ID, if applicable.
- Unique sample identifier.
- Sample type(s).
- Quantity sent and received.
- Research project name and/or number if available.
- Shipping conditions (eg, dry ice, room temperature etc)
- Name/Signature of individual receiving the shipment.
- Any discrepancies between the shipping manifest and the actual shipment.
- Any indication that a biospecimen has been compromised (eg record deviations in sample quality upon receipt).

Please refer to Material Transfer Agreements, Packaging and Shipping Procedure (PR2017/107) for more information.

4.16 Tracking significant events

The informatics system should be able to track a biospecimen through significant events from collection through freezing/thawing, processing, storage, distribution, and possible destruction. This includes tracking of amount distributed and amount remaining of partially-used biospecimens. Restocking of returned, unused samples from the researcher, while not recommended because of potential effects of unknown handling on sample quality, if returned to the research biorepository is must also be tracked. Tracking includes cross-referencing multiple, pre-existing, and/or external physical biospecimen identifiers, such as barcodes with non-identifying information. Any data about the sample being compromised should be noted and available to the user.

4.17 Query capability

The research biorepository database must provide full query capability throughout the system.

5.0 Obtaining confidentiality disclosure agreements

Employees of MSH research biorepositories have access to confidential information in the form of patient medical records. Medical information is protected under federal and state privacy laws and under the terms of the consent process. Furthermore, information may also be bound under non-disclosure or confidentiality agreements. Employees with access to this information may not disclose it.

5.1 Confidentiality Disclosure Agreements — Important Elements

MSH requires that all employees engaged as part of the research biorepository to sign a Confidentiality Disclosure Agreement (CDA) in order to protect sensitive and personal information to which the employee may have access. The Confidentiality Disclosure Agreement must contain at least the following elements:

- definition of confidential information
- knowledge of the appropriate relevant procedures and SOPs
- exclusions (if any) from confidential information
- obligations of the employees
- miscellaneous provisions if relevant

5.2 Confidentiality Disclosure Agreements – completion of agreement

Custodians must request for all research biorepository personnel (that will have access to patient/participant or research information) complete a Confidentiality Disclosure Agreement. The Custodian must obtain, in duplicate, a completed (signed, dated and witnessed by a supervisor) Confidentiality Disclosure Agreement prior to the employee being granted any access to sensitive information. The Custodian must ensure that a supervisor or manager has signed and dated the Confidentiality Disclosure Agreement and retain one copy of the signed and witnessed Confidentiality Disclosure Agreement in hardcopy for the research biorepository records. The employee must be provided with a duplicate copy to for their records. Attachment 4 has a template Confidentiality Disclosure Agreement that may be used by all MSH research biorepositories.