Sample DPIA template

This template is an example of how you can record your DPIA process and outcome. It follows the process set out in our DPIA guidance, and should be read alongside that guidance and the Criteria for an acceptable DPIA set out in European guidelines on DPIAs.

You should start to fill out the template at the start of any major project involving the use of personal data, or if you are making a significant change to an existing process. The final outcomes should be integrated back into your project plan.

Submitting controller details

Name of controller	LSHTM
Subject/title of DPO	Drug Resistant TB contacts electronic registry
Name of controller contact /DPO (delete as appropriate)	DPO: Monica Cozzone

Step 1: Identify the need for a DPIA

Explain broadly what project aims to achieve and what type of processing it involves. You may find it helpful to refer or link to other documents, such as a project proposal. Summarise why you identified the need for a DPIA.

The aim of this project is to help improve and harmonise our approach to Multi Drug Resistant (MDR) TB contact management nationally. We aim to describe TB and latent TB infection (LTBI) incidence rates in exposed close contacts of MDR TB up to 2 years after their exposure. We will do this within an observational cohort study whilst entering the contacts onto a contacts registry held at LSHTM. The follow up consists of three monthly reviews during the first year and six monthly reviews during the second year of follow up

Current international guidance on management of MDR TB contacts is limited by poor evidence and practice varies across the UK and across the globe.

The electronic registry tool has been developed in DHIS-2 (District Health Information Systems 2); an open access software platform increasingly widely used internationally which was created by the University of Oslo. This tool was selected as it is used by the World Health Organisation (WHO) for countries to report their TB surveillance data. Our instance of the DHIS-2 application will be housed on an LSHTM secure server and be used solely for this research project. Due to the nature of MDR TB contact identification personal identifying data will be used and recorded within DHIS-2 including names, dates of birth, phone numbers and addresses.

Participant consent and participation will be localised to each participating NHS MDR TB treatment clinic and paper based with forms scanned and sent to the study team at LSHTM. Health Research Authority (HRA) approval has been granted and NHS Local organisation documents will be generated for each study site.

Once MDR TB contacts are identified as having incident TB they will be referred to the NHS services and managed as per standard of care.

This is a novel type of data processing at LSHTM, DHIS-2 has been used extensively in other universities, across governments and the NGO sector but not previously at LSHTM. The data processed is special category personal data specifically health data and includes data on vulnerable individuals.

Step 2: Describe the processing

Describe the nature of the processing: how will you collect, use, store and delete data? What is the source of the data? Will you be sharing data with anyone? You might find it useful to refer to a flow diagram or other way of describing data flows. What types of processing identified as likely high risk are involved?

Clinico-epidemiological data will be captured using a standardised questionnaire at initial contact screening, three, six, nine, twelve, eighteen and twenty-four month follow up points. The questionnaire will include risk factors for developing TB, previous exposure and treatment of TB and latent TB infection (LTBI), current symptoms of TB, results of Chest Radiographs, TB tests and tests for latent infection. We have minimized the data collected to data required solely for this research project, protecting individuals and improving the tool functionality. The source of the data are the participating individuals. No NHS data will be used. There will be no interaction between NHS health information software and our research DHIS2 software.

Data collection

The data will be collected in TB clinics on NHS computers using a web browser interface. All the NHS computers to be used in this project have user authentication and are based within closed TB clinic rooms. Each participating TB clinic will have a TB specialist nurse trained to enter data with a username and password enabling data entry for their clinic. Two factor user authentication will be used. This will require a one-time password to log in. Once the index and their contacts are registered on the DHIS2 server a study identification (ID) number is generated allowing the participants data to be accessed, viewed and followed up with this pseudonymized code. All data will be entered and viewed via a web browser using usernames and passwords. Each user will have permission to access individual patient data from their clinic only.

Levels of data access are organized through the functions: user, user roles and user groups within DHIS2. This allows allocation of tiers of access and functionality to individuals within a team and to groups of users. On joining the research team end users will be provided with training in DHIS2 use, good clinical practice and research ethics. Their start and end date with the project will be recorded and their DHIS2 access will terminate at their end date. All users will log in at the start of their session with a participant and log out at the end of each individual session.

Data upload to the LSHTM secure server will occur on saving the data entry form in the web browser. Once each review 0, 3, 6 month etc with a participant is complete the user finalizes the form. Changes to data entry can be made to forms subsequently by selecting that the form is incomplete. All these changes are logged within the audit trail.

Data security

The MDR TB DHIS2 instance is hosted on its own dedicated virtual server in the School's data centre, which provides the highest levels confidentiality, integrity and availability. The server's filesystem is stored on secure data volumes, which fall into a segregated backup regime already in place for other secure data housed by the School. All data transmission between the client and server is encrypted over a 2048-bit SSL connection. Once received by the server, data is encrypted with a 24-chararacter key as it is written to the application's PostgreSQL database. Server login details, passwords and encryption keys are stored in the School's 'SecureAnyBox' password management system.

Participants Lists

Only the research team at LSHTM will have access to the separately held full list of study participants with their study participant IDs. These will be held in a locked file on a LSHTM computer. We will not be sharing the personal identifying data with anyone outside of the research team. The anonymised data will be shared with the participating NHS research team members and be used to write up the research for publication. We do expect to report the results of the project to participating NHS institutions. We expect only 20 participants over two years at each of the TB clinics participating with an overall total of 500 participants.

Audit Trail

Audit trails capture and trace all data entered into DHIS2. They run automatically and we will provide oversite, reviewing the audit trail functionality, traceability, data processing, storage and security.

Account management

User accounts will be managed by the LSHTM team who will provide oversight, proactively solving end user concerns. Start, end date and level of access for all users will all be managed by the LSHTM team.

Back up

All data collected will be backed up and stored within a secure server on the LSHTM server platform. Back up will occur automatically on data entry to the LSHTM server.

Recording disclosure control measures

When disseminating data back to NHS sites the data will be presented with limited special category data or personal identifying data. Demographics included will be age and sex but no location data will be included. The data shared will be documented in an audit trail.

Describe the scope of the processing: what is the nature of the data, and does it include special category or criminal offence data? How much data will you be collecting and using? How often? How long will you keep it? How many individuals are affected? What geographical area does it cover?

The data is clinico-epidemiological and contains personal identifiers for the purposes of contact tracing and follow up of recruited participants. There is special category data held but no criminal offense data is held.

At each study visit a questionnaire containing 14 main questions will be asked. These include: the date of contact screening, the date of their index case effective MDR TB treatment start, the duration of index case symptoms prior to MDR TB treatment start, their relationship to the index, their proximity to the index (household exposure risks), risk factors for TB, previous TB or LTBI treatment, current symptoms, chest radiograph, sputum tests and LTBI screening test results. This data will be collected in person at baseline, 6,12, 24 months and over the phone at 3, 9 and 18 months.

The data will be held for 10 years after the study closes. The study is expected to run for 2 years at each study site. Each study site has a median of 2-4 index MDR TB cases/year and each index between 2-3 household contacts. We expect a maximum of 20 contacts from each of the 22 NHS institutions over the two year study. We have HRA ethical approval to run in England and Wales and will be applying for an amendment to work in the MDR TB treating centers in Scotland (of which there are three).

Describe the context of the processing: what is the nature of your relationship with the individuals? How much control will they have? Would they expect you to use their data in this way? Do they include children or other vulnerable groups? Are there prior concerns over this type of processing or security flaws? Is it novel in any way? What is the current state of technology in this area? Are there any current issues of public concern that you should factor in? Are you signed up to any approved code of conduct or certification scheme (once any have been approved)?

The study focuses on MDR TB index patients and their exposed household contacts, many of whom will be family members of the index. The participants will be identified through NHS TB clinics by participating trained clinicians. Participation will not affect their clinical care. Household members individually chose to consent and participate, participation of one household member does not force all household members to participate. Participant information sheets and the consent form includes detail on what participation entails, all participants can choose to withdraw consent at any point without reason and no data will continue to be held. Children of all ages, pregnant women and prisoners are eligible to participate. Individuals who do not have capacity to consent are not eligible.

The software platform is used widely across the world in both research, non-governmental organisations and in national health record systems. The software is Data Health Information System 2 (DHIS2) and is an open access platform developed by the University of Oslo. It is under constant development and dissemination with training academies set up worldwide to increase the software's reach. There are no prior concerns of processing or security flaws.

Participants will be able to withdraw their consent at any point and no further data will be collected. Data already held will continue to be kept in the study. All participants will have access to the LSHTM privacy notice for research participants found here.

Describe the purposes of the processing: what do you want to achieve? What is the intended effect on individuals? What are the benefits of the processing – for you, and more broadly?

The data is required for processing due to the nature of the research. All index patients with infectious MDR TB will have exposed their household contacts to infectious risk. The purpose of the study is to contact these contacts, follow them up over time and quantify the risk.

The benefits for our study team are enabling the work at hand. Wider benefits include improving care for this group of exposed well people who may develop TB or latent TB infection. Additionally, the registry will enable standardization of care across the NHS and provide an evidence base to develop a well overdue guideline for care.

Step 3: Consultation process

Consider how to consult with relevant stakeholders: describe when and how you will seek individuals' views – or justify why it's not appropriate to do so. Who else do you need to involve within your organisation? Do you need to ask your processors to assist? Do you plan to consult information security experts, or any other experts?

As part of this study's development, we have engaged with Principal Investigators at each NHS Trust and via the British Thoracic Society. In this forum we have discussed and debated the lack of effective and systematic treatment for this population across the UK. During a Delphi process with these experts, we consulted and developed a best practice guideline around how to manage this population. This research will gather evidence to support this guideline, improve understanding and raise awareness on this topic allowing the results to be disseminated across the UK. There is broad engagement with this project across clinicians and TB specialist nurses in each NHS Trust.

Within LSHTM we are working with IT services to develop this work and with the Open Data Kit research team experienced in using software for research implementation. A co-investigator is Chrissy h. Roberts, the LSHTM lead for the ODK LSHTM project. The IT support comes from his colleague Matthew MacGregor who is instrumental in ODK development at LSHTM. We have not consulted outside information security experts but discussed the project at length with the LSHTM IT security and legal team.

We have Ethical approval in place from the Health Research Authority, through the process of the NHS research Ethics Committee (REC) and through the LSHTM Research Ethics Committee. The study is sponsored by LSHTM.

Step 4: Assess necessity and proportionality

Describe compliance and proportionality measures, in particular: what is your lawful basis for processing? Does the processing actually achieve your purpose? Is there another way to achieve the same outcome? How will you prevent function creep? How will you ensure data quality and data minimisation? What information will you give individuals? How will you help to support their rights? What measures do you take to ensure processors comply? How do you safeguard any international transfers?

The data is required for household contact identification and follow up, without personal data this research study would not be possible. Linking the index MDR TB patient to their contacts is essential the initial recruitment pathway. The household contacts' names, dates of birth, telephone numbers and home addresses are included in this data. Pseudonymization of household contact data will occur once they have been recruited into follow up. Personal identifiers will continue to be held by the study team in a separate list linked to the participants' study identification numbers.

This study is exclusively for the follow up and investigation of MDR TB household contact screening, management and incident disease identification. There is no plan for any other research with this data set. The data collection is predesigned, due to the hierarchical structure of DHIS2 software build additional data cannot be added once data collection has started.

The individuals recruited will be informed of the data held, the function and purpose of holding this data, where it will be held, how it will be accessed and by whom. There medical care and legal rights will not be affected by a decision to participate or not in this work. All data entry will be username and password specific, enabling a careful audit trail of all work. At each follow up they will be contacted by phone or invited for face-to-face review and be given an opportunity to discuss the study and their inclusion.

The data will be collected within the UK and held at the LSHTM server. Not international organisation will have access to the data and the data will not be transferred internationally.

The lawful basis for processing personal data in this study is UK GDPR Article 6(1)(e): processing is necessary for the performance of a task carried out in the public interest and that task has a basis in law. The specific task takes the form of research and the basis in law is LSHTM's Royal Charter, which empowers LSHTM to perform certain functions to operate as a higher education institution. These functions include "promoting ... research... and education in public health and tropical medicine and such other academic subjects as [LSHTM] may consider appropriate".

The lawful condition for processing special category personal data is UK GDPR Article 9(2)(j): processing is necessary for scientific research purposes or statistical purposes, in accordance with Article 89(1) (as supplemented by section 19 of the Data Protection Act 2018 (DPA 2018)) and with a basis in law.

The relevant basis in domestic law is provided by the DPA 2018, Schedule 1, Part 1, Condition 4.

The processing of personal data anticipated by this DPIA is considered by the LSHTM study team to be a reasonable and proportionate method of achieving the research aims pursued. The study shall not collect more personal data than is needed for these purposes, and as stated above, the processing activities are in the public interest.

As required by Article 89(1) of the UK GDPR and section 19 of the DPA 2018, the LSHTM study team shall ensure that appropriate safeguards are put in place to protect individuals and shall ensure that the specific restrictions set out in these two provisions are complied with. In particular:

- technical and organisational measures shall be implemented to ensure data minimisation – including pseudonymisation, where possible, to make it more difficult to link the personal data back to specific individuals (this shall also act as a security measure during transfer/storage);

- anonymised data shall be used for analysis, publication and further sharing with other internal or external research groups;

- the LSHTM study team do not consider that the processing activities contemplated are inherently likely to cause substantial damage or distress to individuals;

- LSHTM shall only use the data to implement measures or make decisions in relation to individual research participants where this is necessary for the purposes of the study including, but not limited to, participant safety (this is permissible on the basis that the study constitutes 'approved medical research' as defined by section 19(4) of the DPA 2018) and;

- other appropriate safeguards and security measures, as described elsewhere in this DPIA, shall be employed.

In order to ensure that the study meets the highest standards of research ethics, and to satisfy other legal and regulatory requirements outside of data protection, written informed consent (or assent) to participate in the study shall be acquired from participants, as described above.

Step 5: Identify and assess risks

Describe source of risk and nature of potential impact on individuals. Include associated compliance and corporate risks as necessary.	**Likelihood of harm**	**Severity of harm**	**Overall risk**
Note: The risks set out below are a standard set of risks that may arise in the context of any research project involving the processing of personal data. This list of risks is illustrative only and is not exhaustive. Researchers should consider risk from the point of view of their own research project and add to the list below, as appropriate. *Researchers should also rate the risks in the context of their research project, based on likelihood and severity.*	Remote, possible or probable	Minimal, significant or severe	Low, medium or high
Impact of processing on participants: *1. Participants finding data collection overly intrusive* Standard data set collected during health care or research work. Explanation for data collection clear. Options to opt out if this is found to be over intrusive.	possible	Minimal	Low
2. Participants' actual or perceived loss of control over the use of their personal data Individuals whether participants or not may possibly perceive they have reduced control over their data in any health information system or system where their data is routinely collected. In this research study this perception is possible but no different. Given the security features, method of data collection, pseudonymization and structure of DHIS2 within the LSHTM secure server the actual risk is low with minimal harm	possible	Minimal	Low
3. Discrimination of any kind	Remote	Minimal	Low
4. Any reputational damage/emotional distress being caused to participants	Remote	Minimal	Low
5. Physical harm to participants (e.g. from participating in the study data collection activities or in the event of a data breach)	Remote	Minimal	Low
6. Distress caused by a loss of confidentiality	remote	Minimal	Low
7. Re-identification of psuedonymised data	remote	Minimal	Low

8. Other significant economic or social disadvantage to participants	remote	Minimal	Low
Information Security/Personal Data Breach Risks			
9. Loss, unauthorised disclosure of/access to personal data at rest e.g. on tablets, laptops, computers, servers, removable storage or other hardware	remote	Minimal	Low
10. Loss, unauthorised disclosure of/access to personal data in transit.	remote	Minimal	Low
11. Corruption of personal data	remote	Minimal	Low
12. Inaccuracy of personal data	remote	Minimal	Low
13. Unlawful, accidental or improper deletion or alteration of personal data	remote	Minimal	Low
14. Loss of availability of personal data	remote	Minimal	Low
15. Physical security of devices used to process personal data being compromised	remote	Minimal	Low
Other Compliance Risks			
*16. Personal data being stored, used or made available for purposes **other than** archiving purposes in the public interest, scientific or historical research purposes or statistical purposes*	remote	Minimal	Low
17. Proliferation of personal data i.e. personal data being stored in multiple locations, rather than one authoritative source	remote	Minimal	Low
18. Personal data being kept in fully identifiable form longer than necessary	remote	Minimal	Low
Corporate Risks			
19. Reputational risk to LSHTM	remote	Minimal	Low

20. Loss of public trust in LSHTM	remote	Minimal	Low

Step 6: Identify measures to reduce risk

Identify additional measures you could take to reduce or eliminate risks identified as medium or high risk in step 5				
Risk	**Options to reduce or eliminate risk**	**Effect on risk**	**Residual risk**	**Measure approved**
Participants find personal data held for research overly intrusive	This may affect recruitment into the study however once participating we don't foresee this risk as more than a potential risk. To reduce the perceived risk we will provide clarity over what data is held, where it is held, who can see it and what it is used for.	reduced	Low	yes
Participants' perceived loss of control over the use of their personal data	Participants will be invited to review their data held and how it is held at entry to the study and at each follow up visit. Their unique identifier will be given to them and used at each follow up visit.	reduced	low	yes
Distress caused by loss of confidentiality	Careful audit trail in place to identify any loss of fidelity. Regular software review and management. Clear communication with participants if a data breach occurs, including details of where when and how it happened. Ensuring a process of solving the loss of confidentiality and preventing a repeat loss.	reduced	low	yes
Loss, unauthorised disclosure	The small data security risk on entering data through an	reduced	low	yes

of/access to personal data in transit.	NHS computer browser directly onto a LSHTM secure server. We will ensure that all staff entering data are trained to log in and out of the software carefully. Only NHS computers will be used for this process no tablets or personal phones will be used. The secure server will be maintained carefully.			

Step 7: Sign off and record outcomes

Item	Name/position/date	Notes
Measures approved by:	Monica Cozzone, Data Protection Officer, 16/11/2021	Integrate actions back into project plan, with date and responsibility for completion
Residual risks approved by:	N/A	If accepting any residual high risk, consult the ICO before going ahead
DPO advice provided:	Monica Cozzone, Data Protection Officer, 16/11/2021	DPO should advise on compliance, step 6 measures and whether processing can proceed

Summary of DPO advice:

The LSHTM project team are responsible for ensuring the measures identified in Step 6 are integrated into project plan and are implemented prior to the relevant processing activities taking place.

In the event of any material change to or departure from the planned data processing activities, as set out in this DPIA, the DPIA should be updated and resubmitted to the DPO for comment.

The following recommendations are made:

1. LSHTM research staff should have completed the online mandatory data protection training within the last 12 months. It is also highly recommended that they register for the below sessions which also cover data handling:
   - **Producing Data Management Plans for Research Funders** (https://staffbookings.lshtm.ac.uk/index.php/info/2102)
   - **Preparing a DMP for your Research Degree Project** (https://ble.lshtm.ac.uk/course/view.php?id=3518)
   - **Planning and Preparing a Research Bid** (https://staffbookings.lshtm.ac.uk/index.php/info/2094)
   - **Starting a Research grant at LSHTM: things you should know for the first year** (https://staffbookings.lshtm.ac.uk/index.php/info/2097)

2. LSHTM institutional-level policies on data protection, information security, data classification and handling, Data Storage Options - related to Data Classification and Handling Policy and the use of personally-owned devices (i.e. 'bring-your-own-devices’), as well as SOPs issued by LSHTM’s Research Governance and Integrity Office must be adhered to by the LSHTM research team.
3. Ensure that levels of data access have been correctly and appropriately assigned based on "need to know" security principles.
4. Use of firewalls to protect internet connection(s) and the use of up-to-date anti-virus and anti-malware software on all devices used to collect and process personal data in the project.
5. Ensure that there is a mechanism in place for reporting security incidents and Data breaches promptly.
6. Due diligence must be ensured (e.g., use a complex password which you should not tell anyone, be sure no one watches when you enter your password, always log off if you leave your device, etc).

DPO advice accepted or overruled by:	Dave Moore	If overruled, you must explain your reasons

Comments:		
Consultation responses reviewed by:		If your decision departs from individuals’ views, you must explain your reasons
Comments:		
This DPIA will kept under review by:		The DPO should also review ongoing compliance with DPIA