

CHEME Data Protection Impact Assessment

This document records the CHEME Data Protection Impact Assessment (DPIA) process and outcome. It follows the process set out by the Information Commissioner’s Office (ICO) DPIA guidance, and should be read alongside that guidance and the [Criteria for an acceptable DPIA](#) set out in European guidelines on DPIAs.

The steps in the document should be considered at the start of any major project involving the use of personal data, or if making a significant change to an existing process. The outcomes should be integrated back into the project plan.

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 Centre for Health Economics and Medicines Evaluation (CHEME), within the School of Healthcare Sciences, Bangor University undertakes health economic evaluation activities in association with clinical and research departments across the UK. This often involves processing pseudoanonymised patient-level data held by hospitals, GP practices and other academic units. The DPIA is required to ensure we process data in a manner consistent with the General Data Protection Regulations (GDPR). To comply, CHEME must handle personal data using systems and procedures designed and built with consideration of the principles of GDPR. This means we must provide safeguards to protect data (for example, using pseudonymisation or full anonymisation where appropriate), and use the highest-possible privacy settings by default, so that the data is not available publicly without explicit, informed consent, and cannot be used to identify a subject without additional information stored separately. No personal data may be processed unless it is done under a lawful basis specified by the regulation or unless the data controller or processor has received an unambiguous and individualized affirmation of consent from the data subject. For clinical trials, the legal basis is one of “public interest” / “research”.

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?

Data will be pseudoanonymised at source so personal identifiers such as NHS number, patient name and address will be removed at source and replaced with the patient trial randomisation number. Source can include for example hospitals providing Patient-linked information costing (PLICS) data, other academic units providing trial database data and NHS Digital providing HES data. CHEME will not retain any means of identifying patients (e.g. lists of patient names, NHS numbers and addresses).

Data will be collected in encrypted files from source-approved sites such as vocal and sharepoint and stored in encrypted files on either the Bangor U-drive or an encrypted laptop or a flash drive. Storage on the Bangor U-drive will be using restricted-access folders available only to the project researchers whereas encrypted laptops and flash drives holding encrypted data will be stored in locked offices.

Relevant data will be extracted and combined with other pseudoanonymised patient-level data to generate a profile of patient visits to inpatient, outpatient, GP and A&E. These are working files and retained by CHEME in restricted-access folders on the U-drive. They will be used to generate aggregated data for publication in peer-reviewed journals

Source data files such as HES or PLICS data files will be deleted if on the Bangor U-drive and a certificate of destruction obtained from Bangor IT after 3 months. If on an encrypted laptop or flash drive, the data will be shredded using software such as Active Killdisc and a certificate of deletion provided. Source files will only be kept for the time it takes to extract resource use data.

Since the data handled is pseudoanonymised and the source data files are held for a minimal period, the risks are considered minimal in this assessment.

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?

Data will be from across the UK and generally relate to healthcare resource use data such as inpatient stays in clinical trials but may from time-to-time include special category data such as date of death. Clinical trials will range in size from fewer than 10 to as many as 10,000 individuals and may cover anything from six months to ten years depending on what the trial is studying. Data collections are usually specific to a trial and may include a data request halfway through the trial but will always need a final data download at the end of the trial. Datasets will be kept for a minimum period and in any case no more than 12 months; however, information extracted from them relating to healthcare resource use will be retained for longer periods (up to 15 years). Patient personal details such as name, address and NHS number will not be stored alongside these records at any point in the data lifecycle at CHEME.

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)?

CHEME does not have any relationship with the patients and our sole purpose is to generate aggregated health economic data for clinical trials and to report the findings in peer-reviewed journals such as The Lancet or Value in Health either separately or as part of the clinical trial. Patient groups can include children and other vulnerable groups; however, all patients will have read the clinical trial patient information sheet and signed the trial consent form allowing us to process the data. This of course does not affect their right to withdraw from the trial at any time of their choosing and if this is the case, their randomisation details are removed from the randomisation list and their pseudoanonymised data will not be requested. As it is patient-level data, there are many prior concerns over the way the data is processed, therefore, CHEME is in the process of completing its Data Security and Protection (DSP) Toolkit to allow us to measure and publish our performance against the National Data Guardian's data security standards. Until this is published, CHEME addresses these concerns using a system-level security policy (SLSP) which outlines how patient-personal pseudoanonymised data will be handled in a secure manner.

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?

CHEME carries out health economic analysis on new clinical trial interventions to see if they are cost effective. The main method used—cost–utility analysis—considers the quality of life someone will experience as well as the extra life they will gain, as a result of intervening in a particular way. The health benefits are expressed as Quality Adjusted Life Years (QALYs) – or years of good health in lay terms. In general, interventions costing less than £20 000 per QALY are considered by NICE to be cost-effective and represent effective use of resources in the NHS. Effective use of resources is fundamental to enable health and social care providers to deliver and sustain high quality services for people. Without a health economic analysis this could not be achieved.

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?

Patients are consulted right at the start of their journey within a clinical trial. They are provided with a patient information sheet which outlines why their data is needed and how it will be processed. They are also provided with a patient consent form so they can give their permission for their data to be used as well as provided with contact details should they wish to withdraw their consent at any time. At CHEME, we never underestimate the importance of data security and always consult with our partners to ensure the pseudoanonymised data is kept in a secure manner. This means consulting with people providing the data (e.g. universities, hospitals and government bodies such as NHS Digital) and including our own in-house experts on data security to work out the best practical means of receiving, processing and destroying the data in a way consistent with current UK data protection legislation.

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?

Lawful basis is underpinned by informed patient consent to permit receipt, processing and release of data as well as public interest and research. Processing of the data at the patient level achieves the purpose because resource use costs can be derived from resource usage and aggregated into a mean with confidence intervals. Other methods can be deployed, e.g. patient questionnaires; however, these put the onus of reporting down to the patient, themselves, who might not remember all their GP, hospital and outpatient visits in a 12 month period. Function creep is prevented by ensuring no one outside the clinical trial has access to the pseudoanonymised patient-level data and only individuals working within the trial are authorized to access it. Data quality is a task associated with the hospitals providing the data and does not fall into our remit at CHEME; however, checks will be made for example to prevent double counting of hospital stays. Individuals will be given information about their role in the clinical trial in the patient information sheet provided by the clinical trials unit and will be kept updated on the trial by other means such as newsletters, websites and

discussions with their health care professionals. At CHEME, we act outside of the clinical trials unit. We do not contact the patients themselves nor do we store any of their personal information.

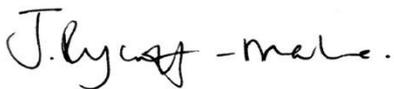
Step 5: Identify and assess risks

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
1. Laptop gets left on a train/in a shop/on a park bench etc	remote	severe	medium
2. Excel file containing data is seen on computer or laptop screen by someone outside of study (e.g. looking over shoulder)	possible	significant	medium
3. Excel file containing data is illegally downloaded from Bangor U-drive by international / corporate / business intelligence hackers	remote	severe	medium
4. Excel file containing data is accidentally downloaded by university students	possible	significant	medium
5. Member of staff wants to access medical record of a patient on a trial	remote	significant	medium
6. Imposters phoning members of the trial study team and asking for information on a patient	possible	severe	high
7. E-mail containing patient data is accidentally forwarded on to someone outside of the trial	possible	severe	high
8. Some other identifier getting into the dataset, e.g. patient postcode, and this being used.	possible	significant	medium

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
1.	Use of encrypted laptops and encrypted folders. No patient identifiers in datasets.	Eliminated	Low	Yes
2.	No patient identifiers in datasets. Persons individualised by randomisation number. Key to randomisation number not held on site.	Eliminated	Low	Yes
3.	No patient identifiers in datasets. Persons individualised by randomisation number. Key to randomisation number not held on site. Original files removed from U-drive or laptop.	Eliminated	Low	Yes
4.	No patient identifiers in datasets. Persons individualised by randomisation number. Key to randomisation number not held on site.	Eliminated	Low	Yes
5.	No patient identifiers in datasets. Persons individualised by randomisation number. Key to randomisation number not held on site.	Eliminated	Low	Yes
6.	Internal policy of not giving patient data out over phone. No patient identifiers in datasets.	Eliminated	Low	Yes
7.	No patient identifiers in datasets. Persons individualised by randomisation number. Key to randomisation number not held on site.	Eliminated	Low	Yes
8.	Extracted datasets will be categorised, e.g. postcodes will be changed to their ranking in the index of deprivation. Original datasets will be deleted	Eliminated	Low	Yes

Step 7: Sign off and record outcomes

Item	Name/date	Notes
Measures approved by:	Colin Ridyard 15 th August 2018 	Integrate actions back into project plan, with date and responsibility for completion
Residual risks approved by:		If accepting any residual high risk, consult the ICO before going ahead
DPO advice provided:	Gwenan Hine 14 th August, 2018 	DPO should advise on compliance, step 6 measures and whether processing can proceed
<p>Summary of DPO advice:</p> <p>Discussions have occurred prior to the DPIA being completed, all advice followed. Section 6 measures are acceptable and have been approved between the DPO and CHEME.</p>		
DPO advice accepted or overruled by:		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:	Colin Ridyard 15 th August 2018 	The DPO should also review ongoing compliance with DPIA