Data Protection Impact Assessment (DPIA)			
https://www.keele.ac.uk/informationgovernance/fortheuniversity/dataprotection/			
Title of Project	National Musculoskeletal Community and Primary Care Audit (National MSK Audit) and Research Database (MSK Research Database)	Your Reference Number (if appl)	Click here to enter text.
Owner (Dept)	Keele Clinical Trials Unit	DPIA conducted by (your name)	Clare Thompson
Date	12/09/2023	DPIA No.(office use only)	DPIA-23-016 updated (v2)

Mandatory grounds to conduct a DPIA	Yes / No?
A1. Will the project using systematic and extensive profiling to make significant decisions about people?	No
A2. Will the project process special category (sensitive) or criminal offence data?	Yes
A3. Will the project systematically monitor publicly accessible places on a large scale (e.g. CCTV)?	No
A4. Are you using new technologies e.g. biometrics, genetic, facial recognition or a major new piece of software?	No
A5. Will the project use profiling or special category (sensitive) data or criminal offence data to decide on access to services, opportunity or benefit?	No
A6. Will the project combine, compare or match data from multiple sources?	Yes
A7. Will the project process personal data without providing a privacy notice directly to the individual ('invisible processing')?	No
A8. Will the project process personal data in a way which involves tracking individuals' online or offline location or behaviour?	No
A9. Will the project process children's personal data for profiling or automated decision-making or for marketing purposes, or offer online services directly to them?	No
A10. Will the project process data that might endanger the individual's physical health or safety in the event of a security breach?	No
If you've honestly answered YES to any of the questions A1 – A10 then it is a legal requirement that you conduct a DPIA – move to Step 1 on page 2	

Advisory Grounds to conduct a DPIA	Yes / No?
B1. Will the project involve **large scale processing** of personal data?	Yes
B1. Will the project involve **profiling** or **monitoring or automatic decision making**?	No
B3. Does the project involve **Special category (sensitive data) or criminal offence data** or the use of the personal data of **vulnerable individuals (including children)**?	Yes
If you've honestly answered YES to any of the questions B1 – B3 then it is strongly recommended that you conduct a DPIA – move to Step 1 on page 2. If you decide not to complete a DPIA even though you've answered Yes in section B then please complete the 'Step A' box on page 2 and then email this form as per instructions in the box below	
If you've honestly answered 'No' to ALL the above question you do not need to conduct a DPIA (although please feel free to do so if you think it would be useful) *PLEASE EMAIL THIS FORM TO: dpo@keele.ac.uk [Please put 'DPIA' in the email subject header]*	

Step 1: Identify the need for a PIA

1.1 Summarise why the need for a DPIA was identified (this can draw on your answers to the screening questions).

As the project will be combining, comparing, and matching data from multiple sources it is felt that a DPIA is required. The project will be linking participants questionnaire answers and MSK treatment and processes of care data from electronic health records (EHR). Informed consent to participate and conduct data linkage will be obtained from individuals to do this.
The project will involve potentially inviting participants who are classed as vulnerable including the elderly and those with mental health conditions.

1.2 Explain broadly what the project aims to achieve, and what type of processing it involves.

The project aims –

- To provide a secure national MSK research database for participating MSK Providers to upload their routinely collected data to for audit and analysis.
- To develop a standard dashboard and reporting system that supports an ongoing automated process to analyse and present the national audit data for quality improvement purposes.

The methods of collecting data will be –

A) Patient survey data – patient reported data capture using either the service's existing electronic data collection platform or a third-party platform commissioned by the study (Netsolving).
B) Primary care electronic health record (EHR) data linkage. EHR data will be captured in NHS systems, e.g., EMIS and SystmOne. All primary care (First Contact Practitioner (FCP)) patients who consent for data to be used in research and data linkage will have data extracted through the FCP template and additional primary care medical record search.
C) Organisational data will be captured for MSK/FCP services at baseline and 6 monthly using Microsoft forms.

1.3 You may find it helpful to link to other relevant documents related to the project, for example a project proposal. (identify other documents here)

National MSK Audit Study Protocol
National MSK Audit Patient Information Sheet
National MSK Audit Privacy Statement

Step A: If you've answered Yes to any B question on page 1 but are not conducting a DPIA, please explain why here (ignore this Step is you are conducting a DPIA)

Click here to enter text.

Step 2: Describe the Processing

2.1 Describe the nature of the processing:

a) how will you collect, use, store and delete data?	Patients will provide informed consent to have their data included in the database. Electronic patient data will be acquired, anonymised, transferred and stored in accordance with Data Protection, GDPR, NHS Information Governance and GCP principles.

	Data will be transferred by service providers to Keele University and initially held within a secure Microsoft SharePoint folder. Direct identifiers such as name, email, telephone contact will be collected by participating Providers using their electronic data capture platform as part of usual care, but these direct patient identifiers will not be shared as part of data exports to Keele University or to the WMSDE. The dataset will be cleaned/checked by the Keele Team and then all of the data will be transferred to the West Midlands Secure Data Environment (WMSDE). Once the data is transferred to WMSDE, the personal data at Keele will be deleted (within 6-months of data transfer). Data within the WMSDE will be held in a secure junction zone, where only the Data Custodian and assigned personnel will have access to the data. No data will be shared outside the environment area and processes will be put in place to ensure that only relevant individuals will have access to personal data and that this is kept separately to the anonymised research database. The database will be held in a secure WMSDE environment and will comply with ISO 27001, ISO 27002, ISO 27017 and ISO 27018. Indirect identifiers Where the database contains indirect identifiers including, NHS number, postcode, date of birth, age, gender, ethnicity, and diagnoses including pain site and diseases etc risk will be managed proportionately when providing access to any data that might, alone or through combination, lead to the identification of an individual. Data Controllers will follow a use-based access control for the purpose of audit, quality checks and reports. The Secure Data Environment (SDE) will be installed on the Microsoft Azure platform and will have the backup and recovery tools provided by Microsoft to protect data and installations. A comprehensive audit trail is in place for the system, and the datasets record these footprints: ● who has accessed the system and when ● when data items are created and who by ● when data items are edited and who by ● when datasets have been browsed, or information (with correct permissions) has been accessed and downloaded; downloads are highly controlled and limited to destinations that are considered to be 'safe settings'. Data Handling and Record Keeping The database will be held in a secure WMSDE owned environment. As with other health data research activities managed by the SDE, this may be held 'on premises' or on cloud or in combination, reflecting the most appropriate platform for the data use in relation to efficiency, function and cost. Regardless of location, the same standards of data security will be required, and will comply with the following standard from the

	International Standards Organisation (ISO): ISO 27001: An international specification for information security management. The corresponding code of practice is ISO/IEC 27002. Cloud provision will be in accordance with the UK Cyber Cloud Principles which are outlined here: https://www.ncsc.gov.uk/collection/cloud-security/implemen ting-the-cloud-security-principles. Additionally, it will comply with the following standards from the ISO: ISO 27017: Code of practice for information security controls based on ISO/IEC 27002 for cloud services. ISO 27018: Code of practice for protection of Personally Identifiable Information (PII) in public clouds acting as PII processors. The database platform will comply with the Department of Health Information Governance policies and standards for secure processing of patient healthcare data, as set out in the Data Security and Protection Toolkit Process of pseudonymisation This is a technical process of replacing personal identifiers in a dataset with other values (pseudonyms), from which the identities of individuals cannot be intrinsically inferred. The WMSDE maintains an association between the original value and replacement value. Examples of this process are replacing an NHS number with another allocated random number curated within the SDE. For the Database the allocated number will be generated using a specific encrypted 'salt code' added to this, before the combined data is then encrypted using a SHA2-256 hashing algorithm. Some internal applicants may require access to use the pseudonymised data in order to support research or quality improvement which requires sequential, longitudinal data reports. Process of anonymisation Most applicants will only access anonymised data and will do so within the WMSDE. On receipt of an approved request, the requested data will be extracted from the pseudonymised Research Database and anonymised. The anonymised data will undergo a QC check by the Data Manager at Keele University for quality and accuracy, and to ensure adequate anonymisation of all data fields. Anonymisation means that information that identifies an individual patient has been removed. The intent of anonymisation is to turn data into a form which does not directly identify individuals, and where re-identification is not likely to take place. This is a technical process of replacing personal identifiers in a dataset with other values, from which the identities of individuals cannot be obtained. The SDE does not maintain any association between the original value and the replacement value. Examples of this process are replacing an NHS number with another allocated random number. It is recognised that patients may choose to opt-out after their data has entered the WMSDE. The Unique Identifier will be available in the pseudonymised dataset. This is to enable the re-identification of patients in the eventuality that they withdraw their consent to be included in future studies. Fully anonymised data that has already been made available to academics within the secure data environment for approved analyses will not be able to be re-identified.

	The Data Custodian will oversee decisions on access to the data. Datasets created on demand will be timestamped and made available under contractual arrangements for prespecified time periods in line with the nature of the projects. Most requests, reviews and release documentation will be stored for 10 years to allow audit and scrutiny of decision-making procedures. Data on any deviations/breaches may be kept indefinitely to allow for assessments of corrective and preventative actions. For certain research records, there will have to be a regard to the Public Records Act 1958 which requires organisations to select records for permanent preservation. Records for preservation must be selected in accordance with the guidance contained in the Records Management Code of Practice 2021.
b) What is the source of the data? (e.g. from the data subjects, from UCAS etc)	• Patient survey data – patient reported data capture using either the service's existing electronic data collection platform or a third-party platform commissioned by the Provider/Study (e.g. Netsolving). • Primary care electronic health record (EHR) data linkage. • Organisational data will be captured for MSK/FCP services at baseline and 6 monthly using Microsoft forms.
c) Will you be sharing data with anyone? (You may find it useful to refer to a flow diagram or another way of explaining data flows)	Yes. Data will be securely transferred to WMSDE.
d) What types of processing identified as likely high risks are involved?	Once data is transferred from Keele to WMSDE the personal data held at Keele will be deleted. No data will be shared outside the environment area and only relevant individuals will have access to personal data which will be kept separate to the anonymised research database. Data Controllers will follow a use-based access control for the purpose of audit, quality checks and reports. The Secure Data Environment (SDE) will be installed on the Microsoft Azure platform and will have the backup and recovery tools provided by Microsoft to protect data and installations. A comprehensive audit trail is in place for the system, and the datasets record these footprints: ● who has accessed the system and when ● when data items are created and who by ● when data items are edited and who by ● when datasets have been browsed, or information (with correct permissions) has been accessed and downloaded; downloads are highly controlled and limited to destinations that are considered to be 'safe settings'.

	Data Handling and Record Keeping The database will be held in a secure WMSDE owned environment. As with other health data research activities managed by the SDE, this may be held 'on premises' or on cloud or in combination, reflecting the most appropriate platform for the data use in relation to efficiency, function and cost. Regardless of location, the same standards of data security will be required, and will comply with the following standard from the International Standards Organisation (ISO): ISO 27001: An international specification for information security management. The corresponding code of practice is ISO/IEC 27002. Cloud provision will be in accordance with the UK Cyber Cloud Principles which are outlined here: https://www.ncsc.gov.uk/collection/cloud-security/implemen ting-the-cloud-security-principles. Additionally, it will comply with the following standards from the ISO: ISO 27017: Code of practice for information security controls based on ISO/IEC 27002 for cloud services. ISO 27018: Code of practice for protection of Personally Identifiable Information (PII) in public clouds acting as PII processors. The database platform will comply with the Department of Health Information Governance policies and standards for secure processing of patient healthcare data, as set out in the Data Security and Protection Toolkit. Where the database does contain NHS number, postcode, age, gender and diagnosis, risk will be managed proportionately when providing access to any data that might, alone or through combination lead to the identification of an individual.
2.2 Describe the scope of the processing:	
a) What is the nature of the data, and does it include special category or criminal offence data?	Data includes identifiable data (NHS number) and medical data.
b) How often will the data be processed	Data will be processed throughout the study period.
c) How long will you keep it?	10 years
d) How many individuals are affected (approx.)?	MSK Community Services – aiming to upload data for a minimum of 250 patients per service (aiming for a minimum of 20 services) to complete a baseline and follow up questionnaire, aiming for a minimum of 5,000 patients per 12-month reporting period when all services onboarded. Primary Care FCP Services – aiming to upload data for a minimum of 25 patients from a minimum of 20 FCP providers (aiming for a minimum of 500 patients), per 12-month collection period when all services onboarded.

e) What geographical area does it cover?	National
2.3 Describe the context of the processing:	
a) What is the nature of your relationship with the individuals?	Data controller of subject's personal data
b) How much control will they have? (e.g. can they access their info and amend / check etc?)	Participants will not have access to their data.
c) Would they expect you to use their data in this way?	Yes, information about how the data is to be used will be given to the subjects in the Participant Information Sheet.
d) Do they include children or other vulnerable groups?	Yes Adults over the age of 18 will be invited to participate, including the elderly, those with disability and mental health issues.
e) Is it particularly novel in any way? (new technologies for instance)	No
f) What is the current state of technology in this area?	Not applicable
g) Are there any current issues of public concern or other concerns that you should factor in?	No
h) Is Keele signed up to any approved code of conduct or certification scheme (once any have been approved)?	[DPO: No / None approved]
2.4 Describe the purposes of the processing:	

a) What do you want to achieve?	The aim of processing this data is to develop a secure national MSK research database for participating MSK Providers to upload their routinely collected data to for audit and analysis. The data will measure the quality and effectiveness of care for patients presenting in general practice and community MSK services and allow us to develop a standard dashboard and reporting system to analyse and present the national audit data for quality improvement purposes, aiming to improve the quality and equity of care provision for patients presenting with MSK conditions through enhanced reporting and evaluation of quality data for participating services.
b) What is the intended effect on individuals?	There may not be any immediate benefits for individuals, although some people find it rewarding to take part in health research.
c) What are the benefits of the processing for you and more broadly?	The study will develop our ability to collect/report MSK data at scale, to share data from multiple providers of NHS care and to bring this data together in a secure data warehouse/environment. The collective data will then be used to identity variation in care provision and patient outcomes, and to understand, share and promote best practice.

Step 3: Consultation Process	
3.1 Consider how to consult with relevant stakeholders:	
a) Describe when and how you will seek individuals' views – or justify why it's not appropriate to do so.	We have proposed these processes to our Patient Advisory Group and independent Expert Advisory Board who have provided their advice and guidance.
b) Who else do you need to involve within Keele e.g. Info Security?	DPIA submitted to DPO. Keele CTU will liaise with appropriate technical personnel for the secure electronic storage of data within Keele University and the transfer of data to WMSDE.
c) Do you need to ask any of your Data Processors to assist (who if so)?	Privacy Statement states: - Netsolving - Provider of the British Society of Rheumatology (BSR) electronic patient reported outcome measures (ePROMs) platform Providers of online data collection platforms commissioned by musculoskeletal healthcare providers to capture patient survey data Academic partner for curation of pseudonymised data (National Musculosketal Audit and Research Database) WMSDE – commissioned by Keele to store / host the research database
d) Do you plan to consult information security	No

experts, or any other experts?	
Note : You can use consultation at any stage of the DPIA process.	

Step 4: Assess necessity and proportionality	
4.1 Describe compliance and proportionality measures, in particular:	
1) What is your lawful basis for processing (see website for info)?	GDPR Conditions: a) Article 6(1)(e) The processing is necessary for the performance of a task carried out in the public interest or in the exercise of official duty vested in Keele (as the Data Controller) AND Article 9(2)(j) – scientific research purposes
2) Does the processing actually achieve your purpose?	Yes
3) Is there another way to achieve the same outcome?	No
4) How will you prevent function creep (i.e. using the data for other purposes)?	Strict procedures and protocols are in place which will prevent function creep. The study also has oversight from an independent Advisory Board.
5) How will you ensure data quality? (how do you ensure the data is accurate) and data minimisation (that the processing only involve the least amount of data for the purposes)?	All study documents are designed to capture the minimum amount of data required to meet the purpose of the study. Regular monitoring of the data will take place by a study management group. The study also has oversight from and independent Advisory Board. Data will receive regular data verification check in line with Keele CTU data processing policy.
6) What information will you give individuals? (transparency info)	Please see attached Participant Information Sheet and the Privacy Statement, both of which will be made available to patients as part of the consenting process.
7) How will you help to support their rights?	As described in the Participant Information Sheet a link is provided to access Keele University Information Governance policy.
8) What measures do you take to ensure processors comply?	Data processing, storage and archiving processes will be delivered in line with Keele University policies and Quality Management System.
9) Are there any international transfers and if so how do you safeguard them?	No

Step 5: Identify and assess risk

Risk Key			
Likelihood of harm	**Severity of harm**		
	1 - Minimal	2 - Significant	3 - Severe
1 – Remote	**Low 1**	**Low 4**	**Medium 7**
2 – Possible	**Low 2**	**Medium 5**	**High 8**
3 – Probable	**Medium 3**	**High 6**	**High 9**

Describe the source of risk and nature of potential impact on individuals. Include associated compliance and corporate risks as necessary	**Likelihood of harm** **Remote, Possible or Probable**	**Severity of harm** **Minimal, Significant or Severe**	**Overall Risk** **Low, Medium or High**
Risk: Mishandling of personal data by staff causing a breach.	Remote	Significant	Low
Risk: Unauthorised access to data	Remote	Significant	Low
Risk: NHS number used by staff to request data from GP practices; staff handling special category data and additional risk to this being disclosed accidentally by staff	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 Eliminated, Reduced or Accepted	Residual risk Low, Medium or High	Measure Approved? Yes/No
Click here to enter text.		Choose an item.	Choose an item.	Choose an item.
Click here to enter text.	Click here to enter text.	Choose an item.	Choose an item.	Choose an item.
Click here to enter text.	Click here to enter text.	Choose an item.	Choose an item.	Choose an item.

Step 7: Sign off and record outcomes

It is the Project lead's responsibility to:
*1. review consultation responses; 2. Seek DPO advice; 3. Accept DPO advice **or** get DPIA reviewed by the SIRO (who can overrule DPO advice); and 4. Approve the final agreed measures in step 6 and integrate back into project design.*

Item	Name	Date	Notes
Project Lead			
Consultation Responses reviewed by (usually Project lead)	Clare Thompson	12/09/2023	If your decision departs from individuals' views, you must explain your reasons: Click here to enter text.
DPO			
DPO advice provided	Anne-Marie Long	18/05/202312/09/2023	DPO should advise on compliance, Step 6 measures and whether

			processing can proceed.
Summary of DPO advice	Having considered the DPIA, the risks identified are mitigated by the processes put in place as part of the processing and so overall remains low. Processing can continue with adherence to the outlined procedures and appropriate participant information leaflets and privacy notices / data processing agreements being in place which the PAC team will assist with. Updated 14/09/2023 No change to the processing of the personal data has been undertaken; reference to the Information Governance Toolkit has been updated to reflect its current name DSPT as well as retention of the data included although this was already reflected in the PIS and generally standard research data is kept for this period of time. No change to the risk and processing can continue as previously advised.		
SIRO (where applicable)			
DPO Advice accepted or overruled by SIRO	N/A		If overruled, the SIRO must explain reasons below (If accepting any residual high risk, consult the ICO before going ahead)
Comments	N/A		
Project Lead			
Measures approved by (usually Project lead)	Clare Thompson	14/09/2023	Integrate actions back into project plan, with date and responsibility for completion.
This DPIA will be kept under review by	Clare Thompson – Trial Manager		The DPO should also review ongoing compliance with DPIA

Completed DPIA should be emailed to:

dpo@keele.ac.uk

Please put DPIA in the email subject header

You should keep a copy of the DPIA on file. Be prepared to produce this if required to by the ICO